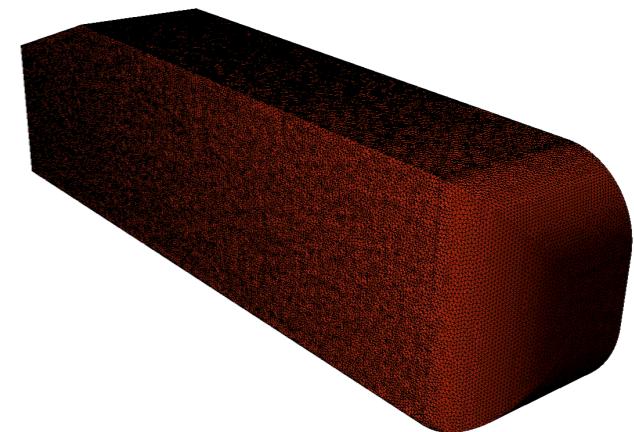
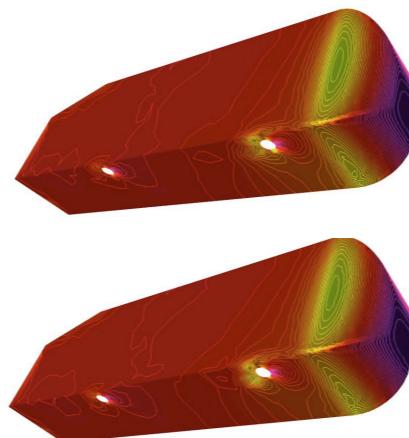
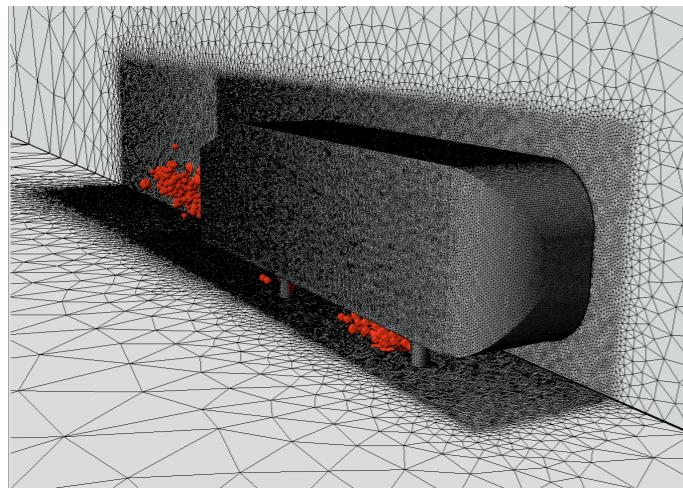


The GNAT method for nonlinear model reduction: discrete optimality, practical implementation, & application to CFD



Kevin Carlberg*, Charbel Farhat[#], Julien Cortial*, David Amsallem[#]

*Sandia National Laboratories

#Stanford University

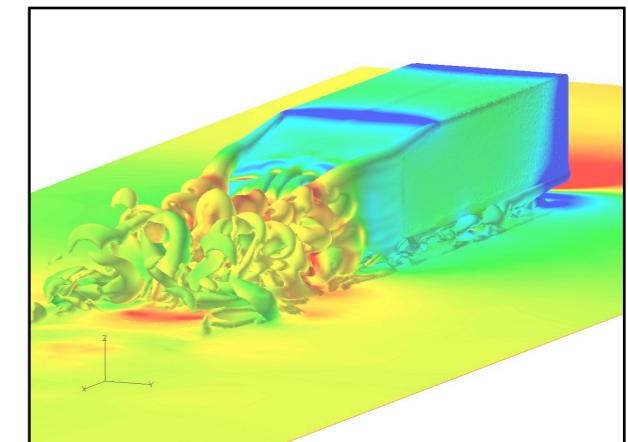
ACDL Seminar

April 17, 2013

Motivation

High-fidelity simulation

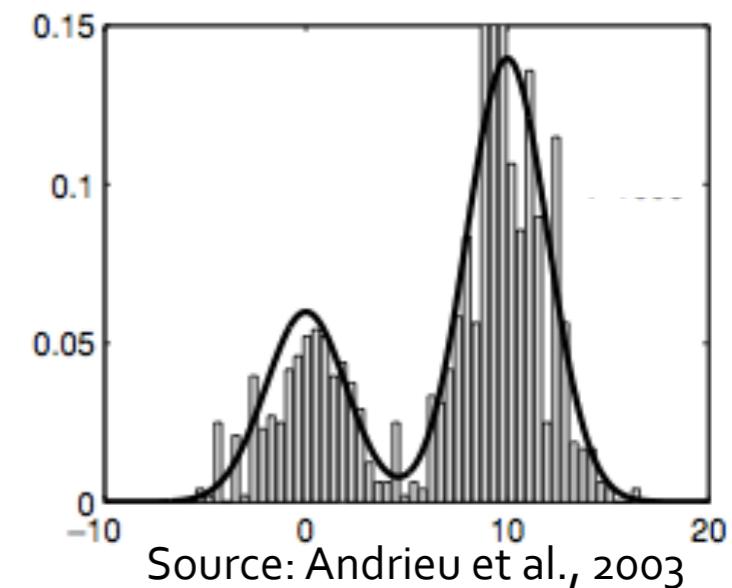
- An indispensable tool
- However, only **a few** simulations are feasible



barrier

Many-query applications

- Require **thousands** of simulations
 - Uncertainty quantification
 - Time-critical design



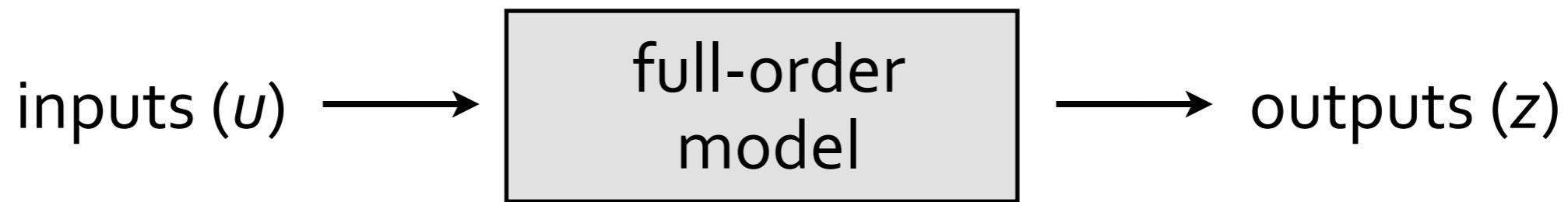
Goal: break barrier

Time-critical design



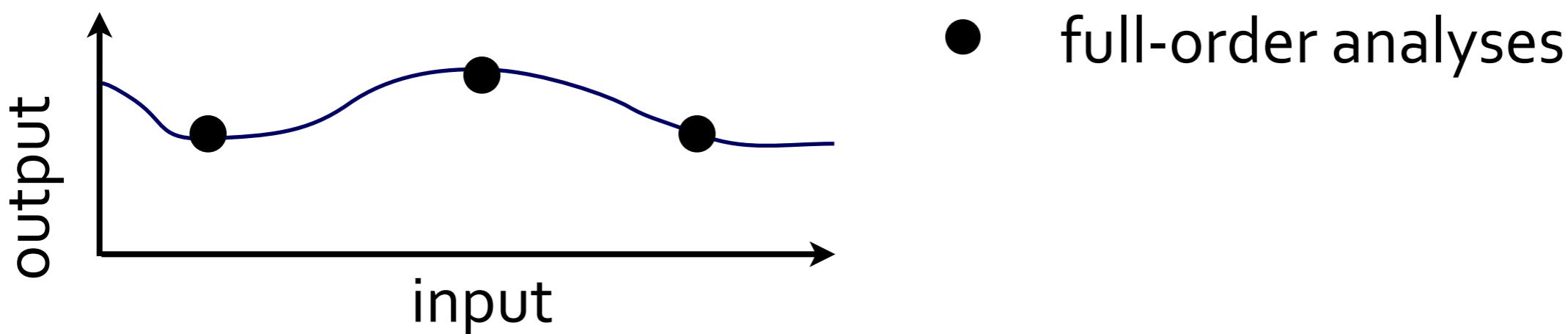
- *Goal:* ~100 detailed aerodynamics simulations per week *during* season
- *Obstacle:* 6 weeks on a supercomputer for one simulation
- *Opportunity:* costly analyses *before* season

Offline/online surrogate strategy



1) Offline

- full-order analyses
- build surrogate model



Offline/online surrogate strategy

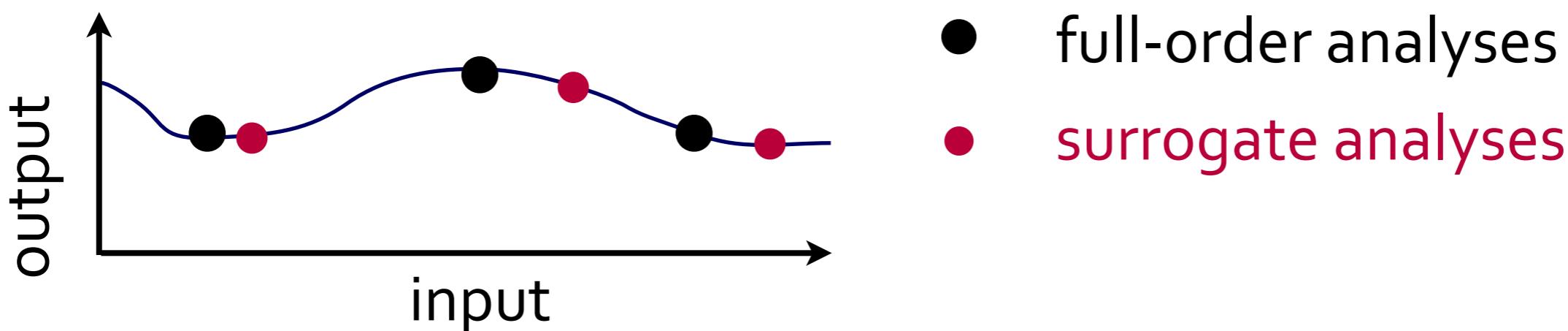


1) Offline

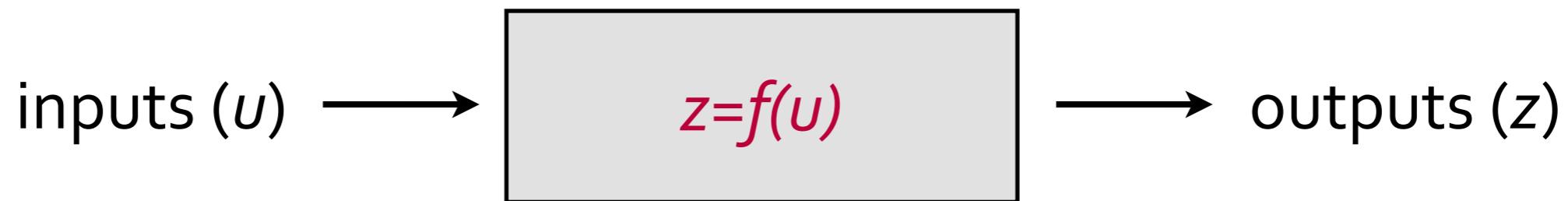
- full-order analyses
- build surrogate model

2) Online

- analysis with **surrogate model**

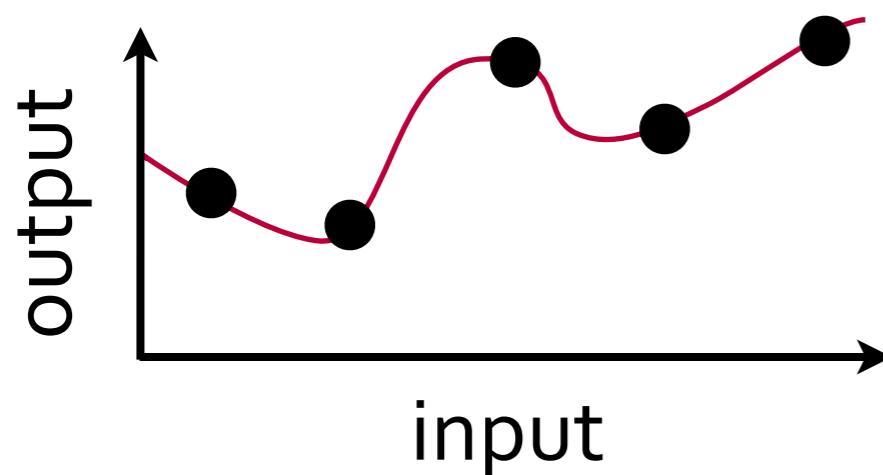


Surrogate-modeling methods



1) Data fits

- model input-output map
 - + black-box implementation
 - blind to physics
 - curse of dimensionality

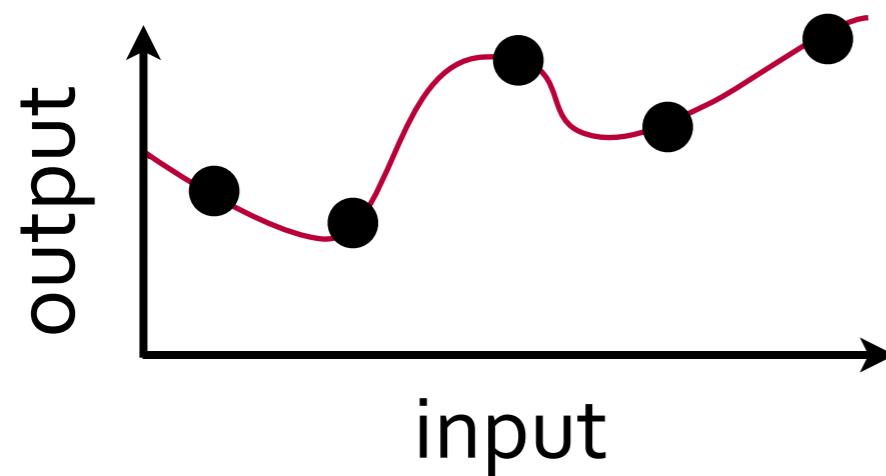


Surrogate-modeling methods



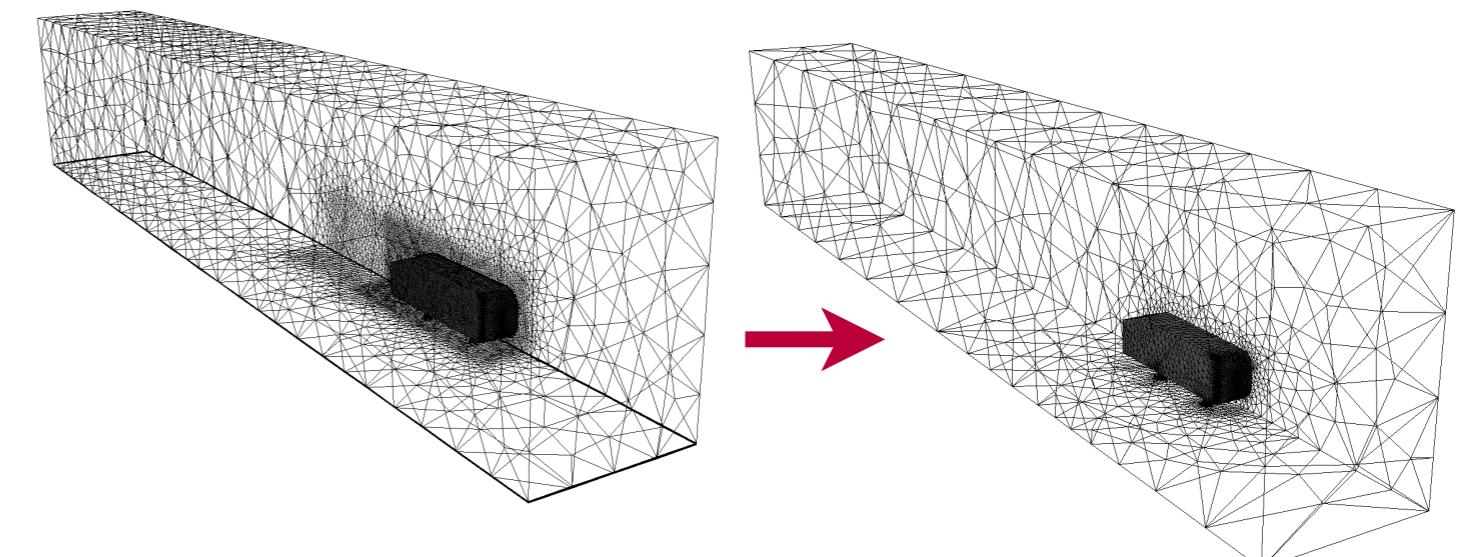
1) Data fits

- model input-output map
 - + black-box implementation
 - blind to physics
 - curse of dimensionality



2) Lower-fidelity model

- coarser mesh, omitted physics,
lower element order
 - + queries physics
 - limited speedup

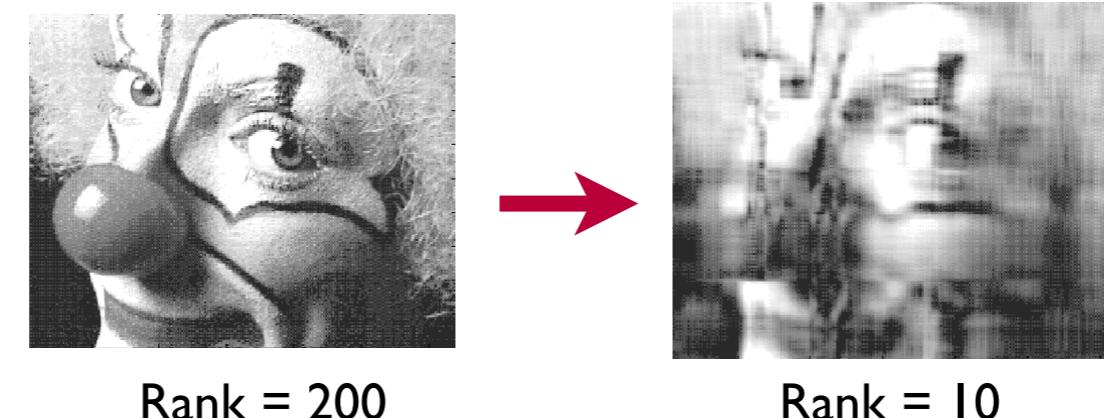


Surrogate-modeling methods



3) Reduced-order modeling

- Extracts *dominant physical behavior* from full-order model
 - + physics computations
 - + can exploit problem structure
 - + potential for high speedups
 - + error bounds
 - intrusive to implement
 - difficult to derive



Reduced-order modeling

- ‘Mature’ for specialized systems
 - linear time-invariant (LTI) systems
[Book: Antoulas, 2005; Willcox & Peraire, 2002; Rowley, 2005]
 - elliptic & parabolic equations
[Book: Patera & Rozza, 2006; Veroy et al., 2003; Grepl et al., 2005]
- Nonlinear systems
 - linearization + LTI method
[Chen & Kang, 2000; Bai, 2002; Rewienski & White, 2003]
 - Proper Orthogonal Decomposition (POD)–Galerkin
[Sirovich, 1987; Lall et al., 2002]
 - approximated POD–Galerkin
[Grepl et al., 2007; Barrault et al., 2007; Chatarantabut & Sorensen, 2010; Galbally et al., 2010]

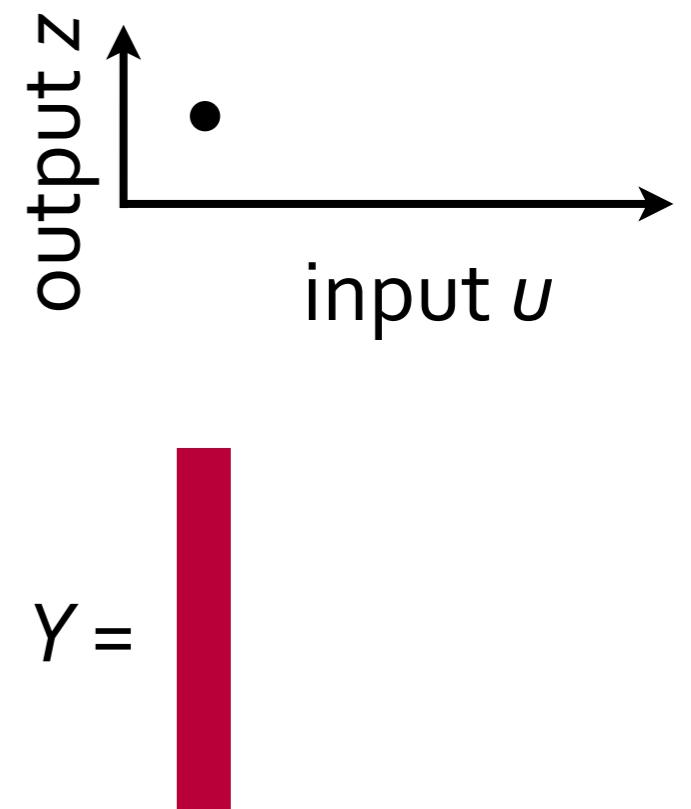
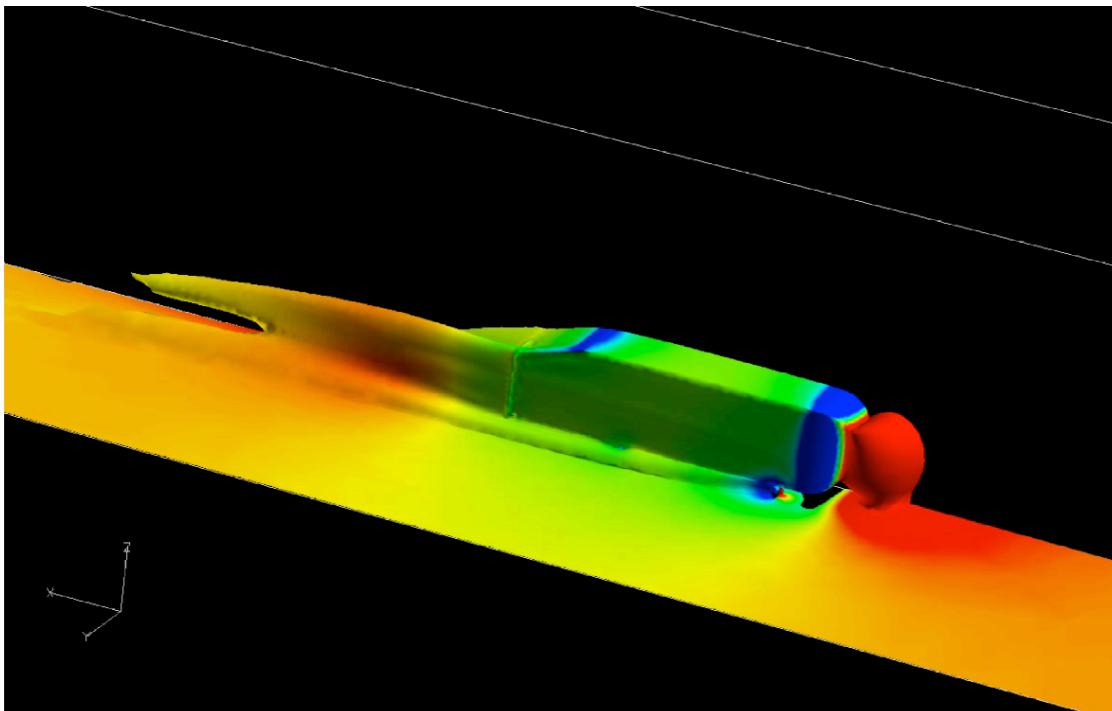
Outline

- Introduction
 - Objectives & background
 - POD–Galerkin
- Gauss–Newton with Approximated Tensors
- Implementation
- Examples

Projection

Offline data collection at training inputs

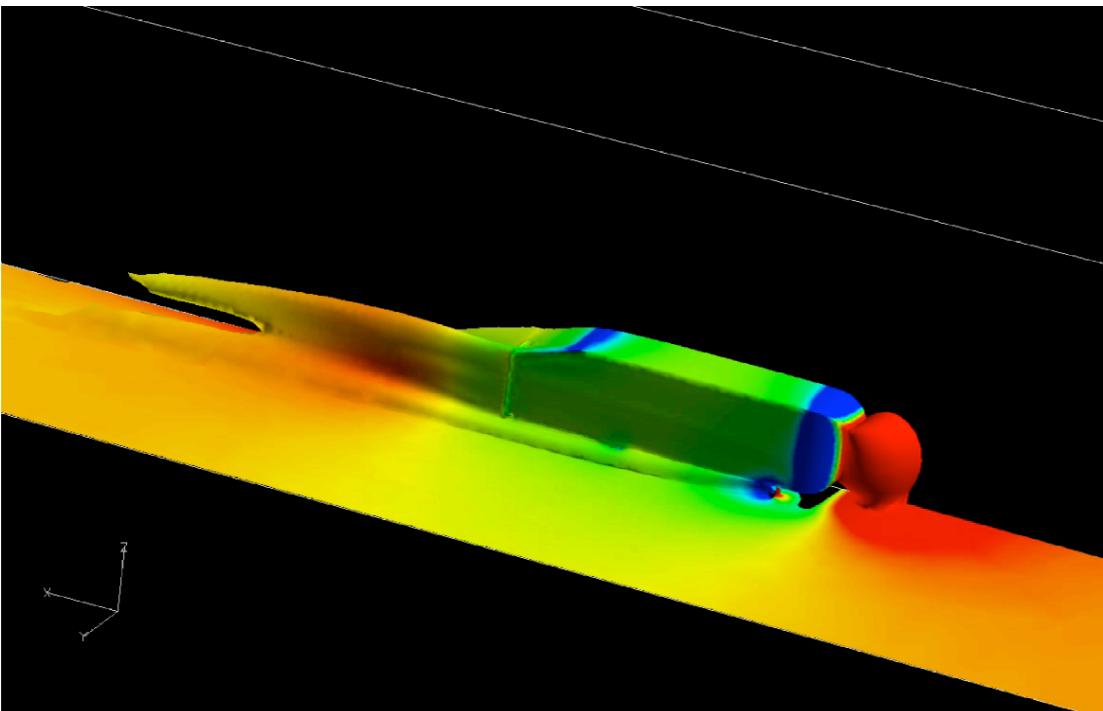
1. Collect snapshots of the state vector



Projection

Offline data collection at training inputs

1. Collect snapshots of the state vector

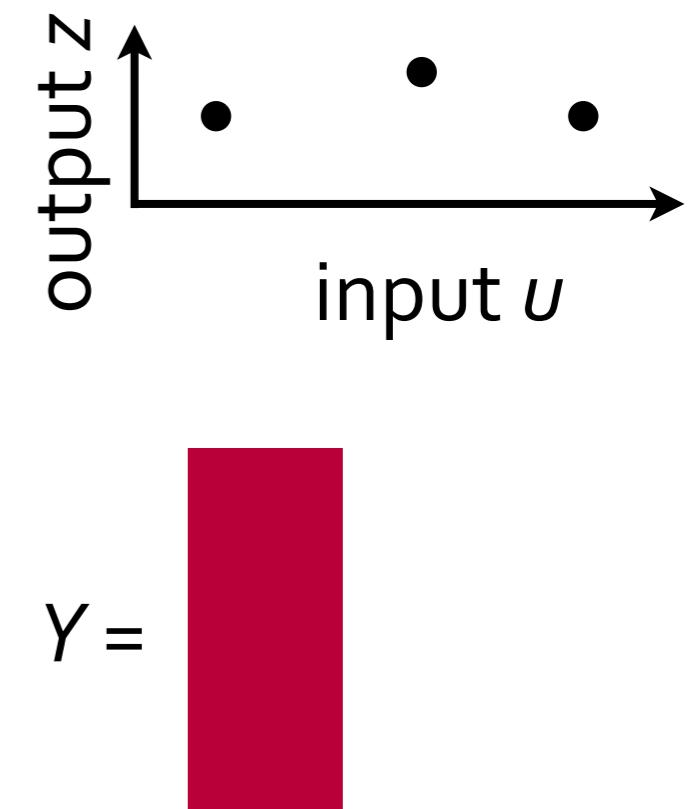


2. Compression

- a. compute singular value decomposition

b. truncate

$$U$$

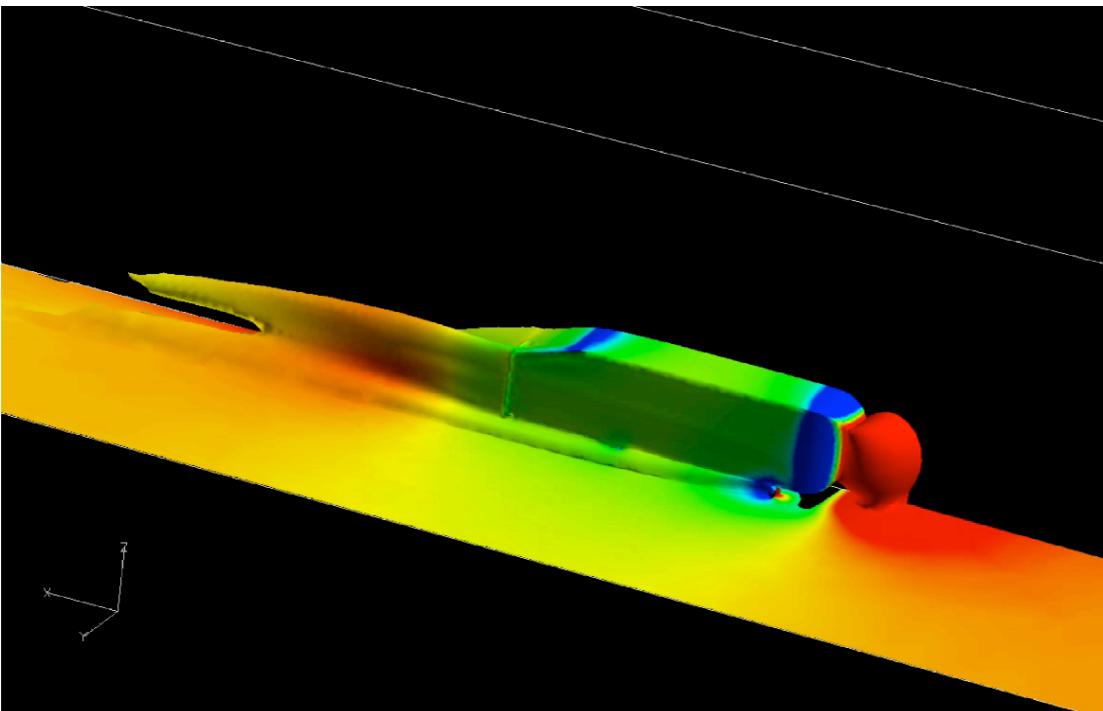


$$Y = U \Sigma V^T$$

Projection

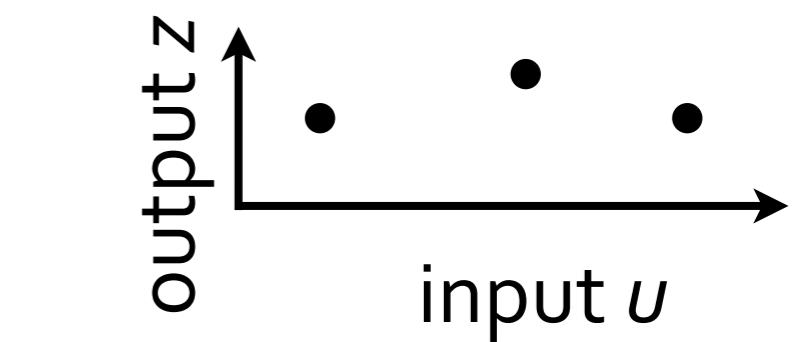
Offline data collection at training inputs

1. Collect snapshots of the state vector



2. Compression

- a. compute singular value decomposition



$$Y =$$

$$Y = U \Sigma V^T$$

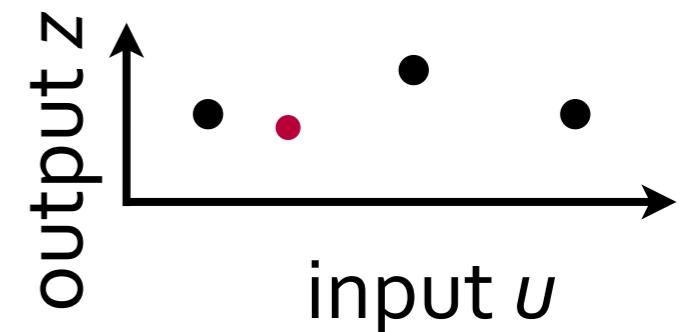
- b. truncate Φ



POD–Galerkin

Online

$$\dot{y} = f(y; t, u)$$



- Galerkin projection

reduce # unknowns

$$y \approx \Phi y_r$$

$$\boxed{\quad} \approx \boxed{\quad} \boxed{\quad}$$

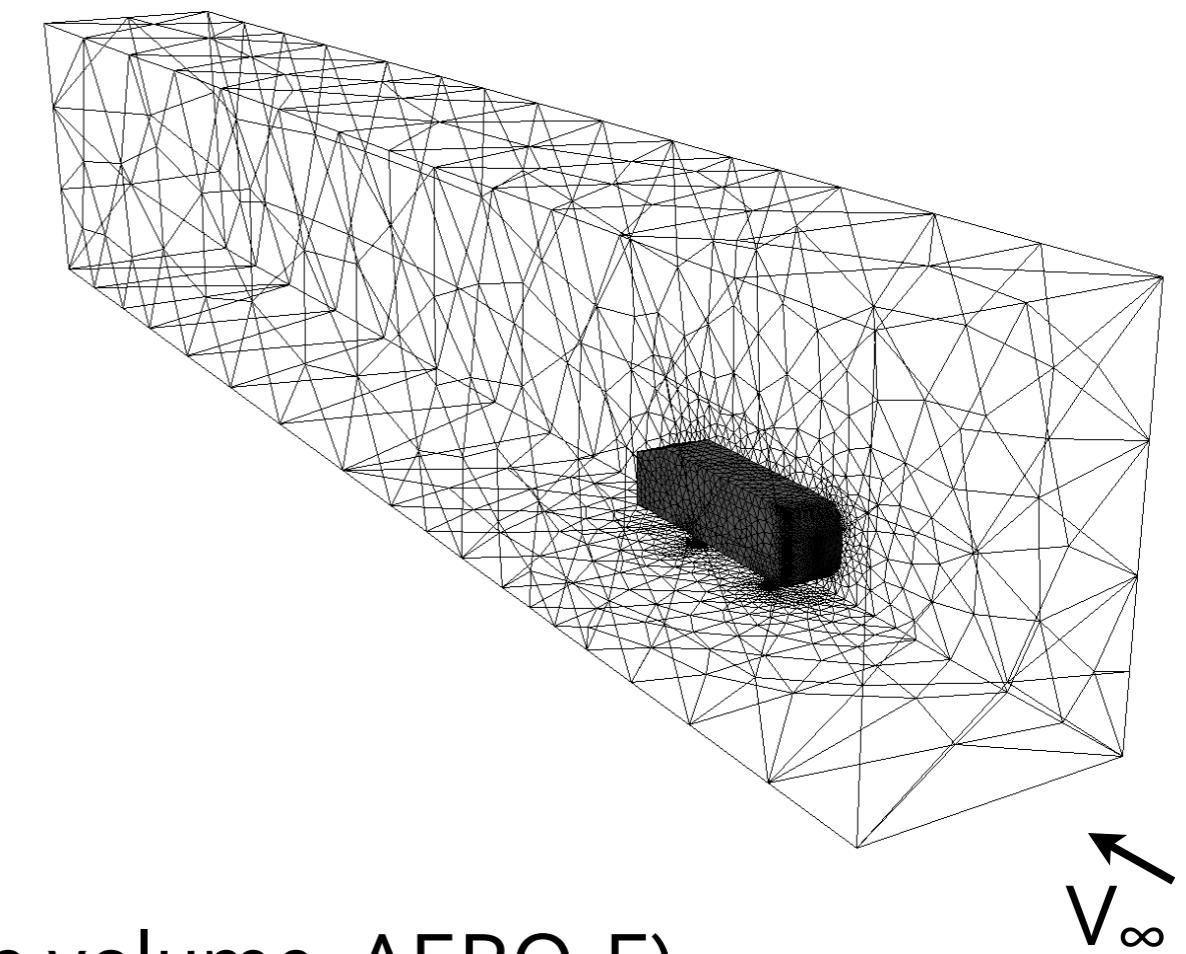
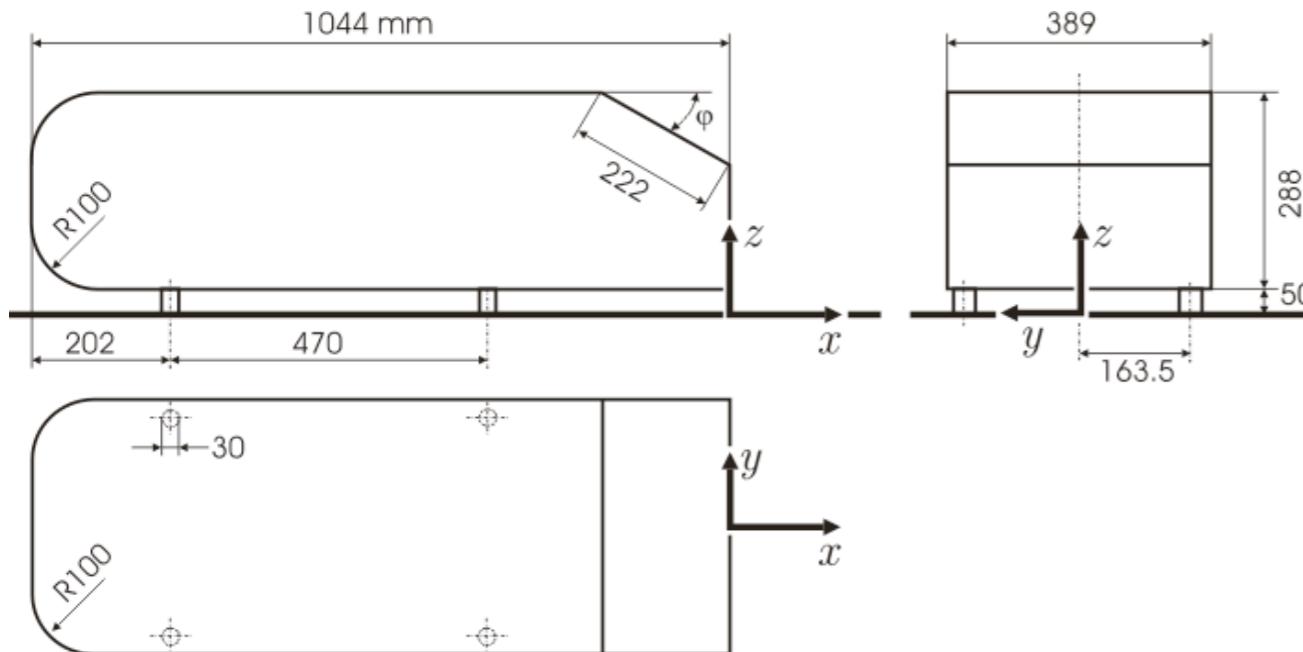
reduce # equations

$$\Phi^T [\dot{y} - f(y; t, u)] = 0$$

$$\boxed{\quad} = 0$$

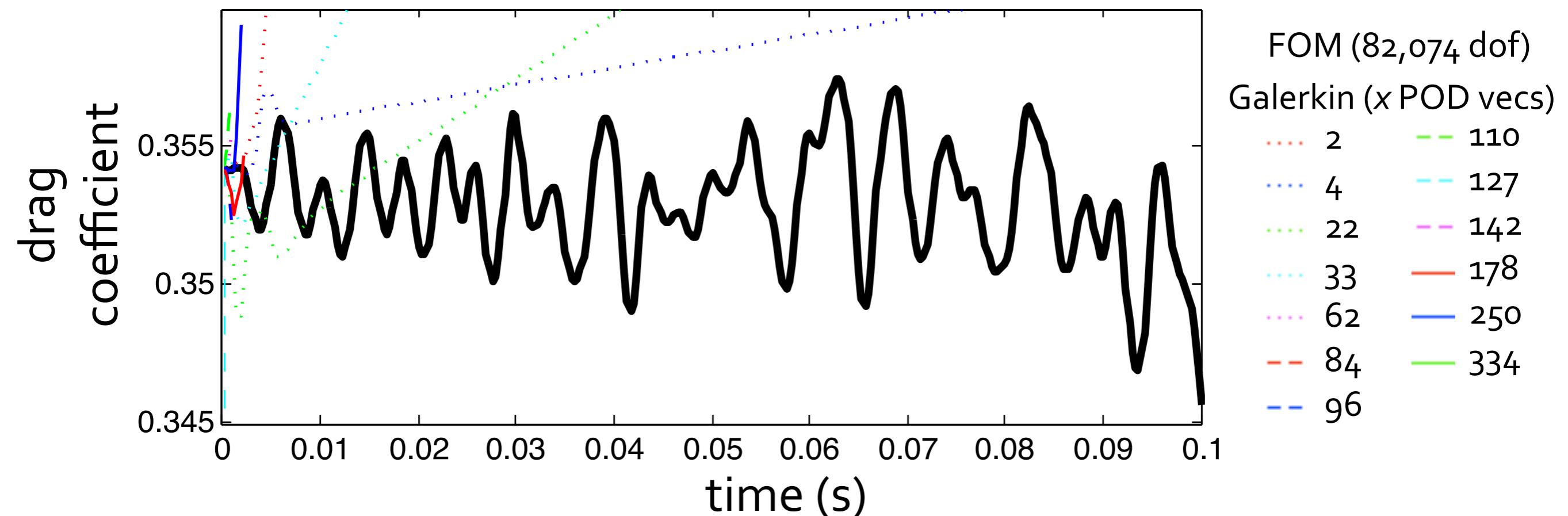
$$\dot{y}_r = \Phi^T f(\Phi y_r; t, u)$$

Benchmark: Ahmed body

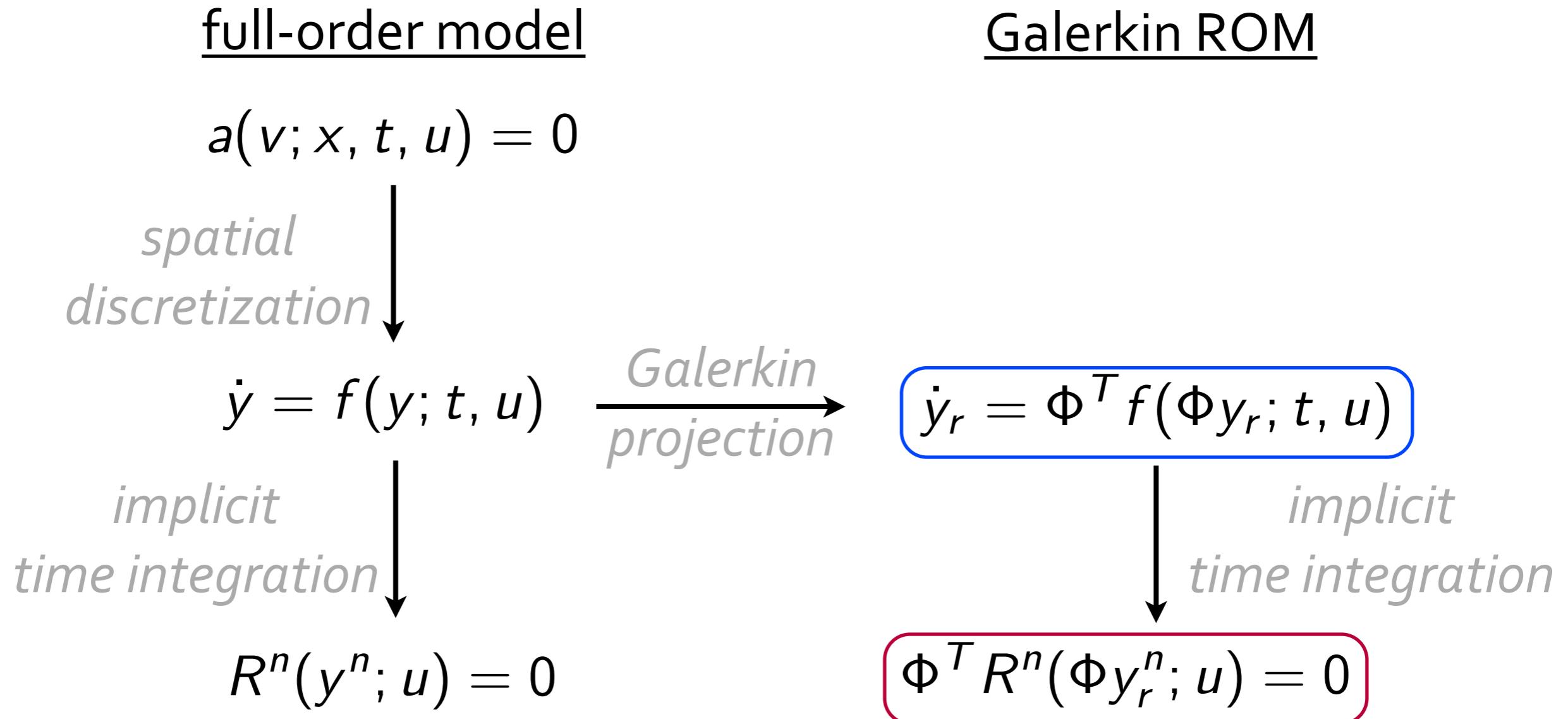


- Compressible Navier-Stokes (finite volume, AERO-F)
 - $\text{Re} = 4.48 \times 10^6$
 - $M_\infty = 0.175$ (134 mph)
 - steady-state initial condition
 - 2nd order flux reconstruction
 - DES turbulence model
 - Spalart–Allmaras RANS
 - $\Delta t = 3 \times 10^{-4}$
 - **82,074** degrees of freedom (dof)

Galerkin ROM (with POD) is unstable



Galerkin ROM is not 'discrete optimal'



+ Semi-discrete optimal: $\Phi \dot{y}_r = \arg \min_{x \in \text{range}(\Phi)} \|\dot{y} - x\|_2$

- Not discrete optimal: $\begin{cases} \Phi \dot{y}_r^n \neq \arg \min_{x \in \text{range}(\Phi)} \|y^n - x\| \\ \Phi y_r^n \neq \arg \min_{x \in \text{range}(\Phi)} \|R^n(x)\| \end{cases}$

GNAT: main idea

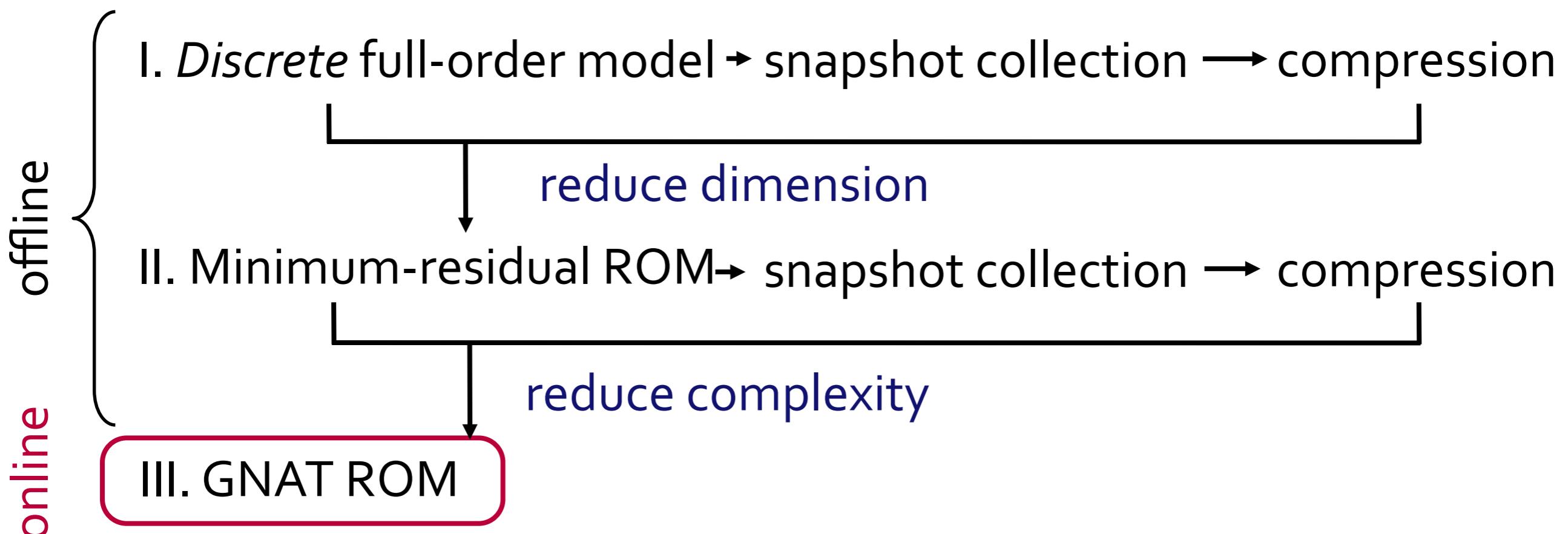
- **Goal:** practical nonlinear model reduction
- **Problem:** POD–Galerkin often does not work
 - lacks *discrete optimality*
- **Idea:** GNAT model reduction
 - *discrete-optimal* approximations
 - effective ‘sample mesh’ implementation
 - works on large-scale problems

Outline

- Introduction
- Gauss–Newton with Approximated Tensors (GNAT) [Carlberg et al., 2011]
 - Approach
 - Error bound
- Implementation
- Examples

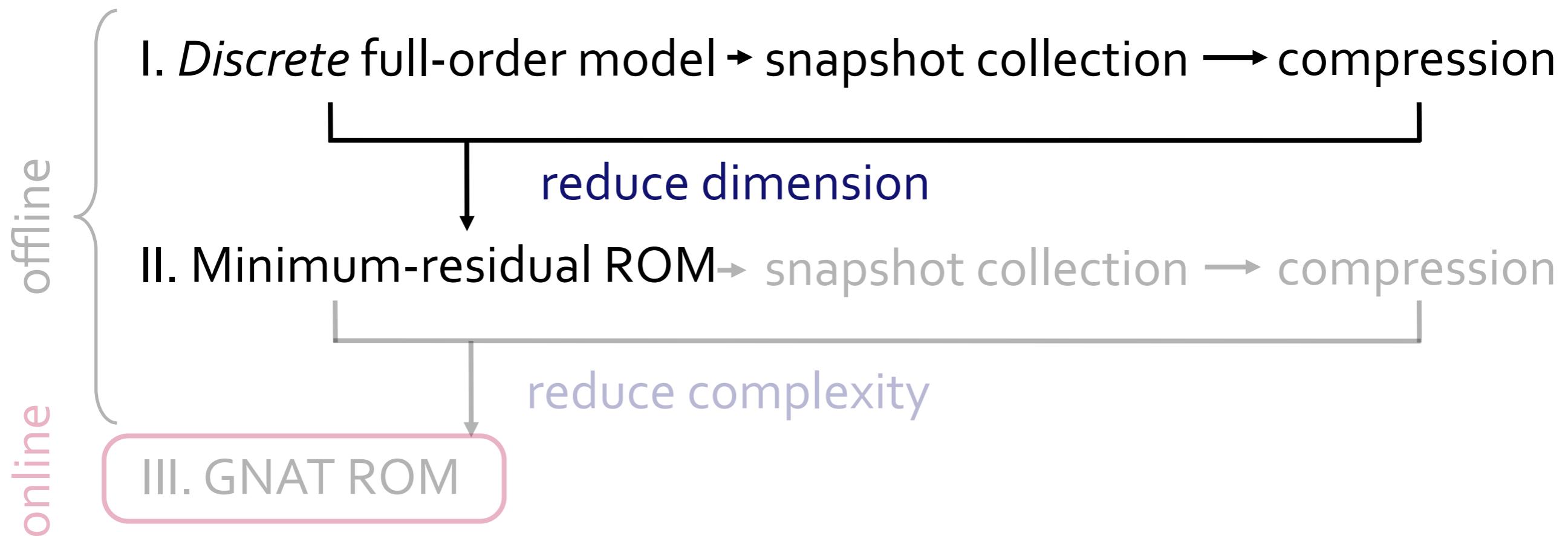
GNAT strategy

- Discrete-optimal approximations



GNAT strategy

- Discrete-optimal approximations: minimize *discrete-system error* over approximation space



Min-residual ROM is discrete optimal

Minimum-residual ROM

full-order model

Galerkin ROM

$$a(v; x, t, u) = 0$$

*spatial
discretization*

not defined

$$\dot{y} = f(y; t, u) \rightarrow$$

$$\dot{y}_r = \Phi^T f(\Phi y_r; t, u)$$

implicit

time integration

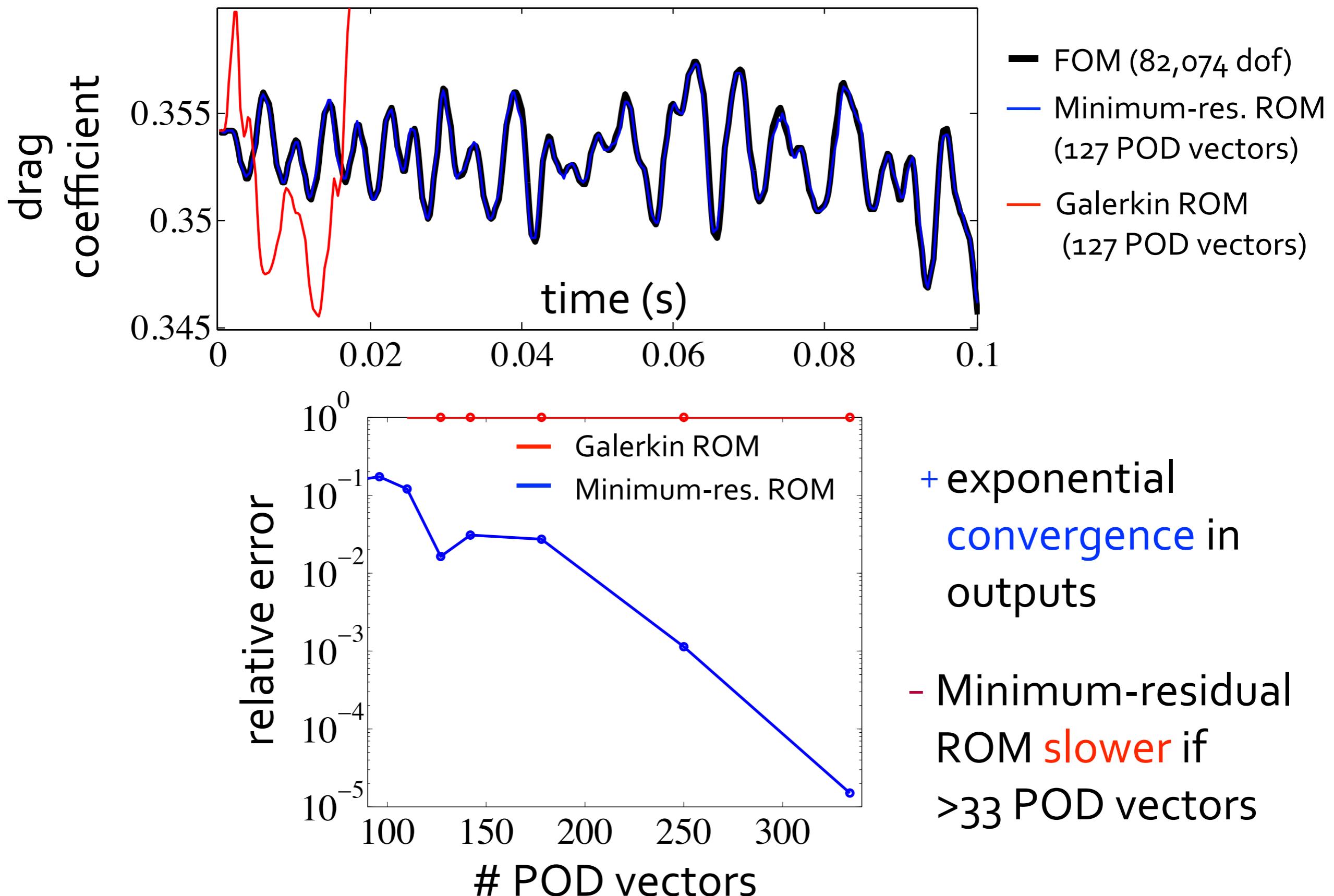
$$y_r^n = \arg \min_x \|R^n(\Phi x; u)\|_2$$

$$\leftarrow R^n(y^n; u) = 0$$

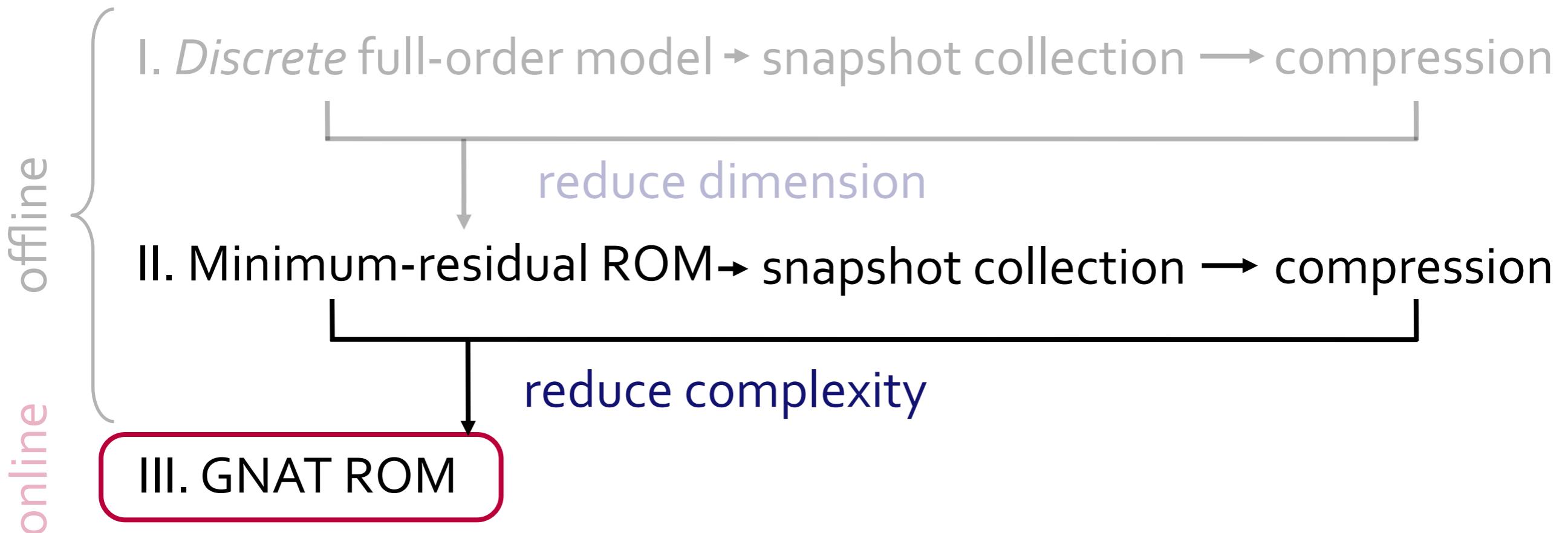
$$\Phi^T R^n(\Phi y_r^n; u) = 0$$

- + Discrete optimal
- Not defined at the semi-discrete level
- Other least-squares ROM methods: Legresley & Alonso, 2001;
Bui-Thanh, *et al.* 2008; Constantine & Wang, 2012

Minimum-residual ROM is accurate



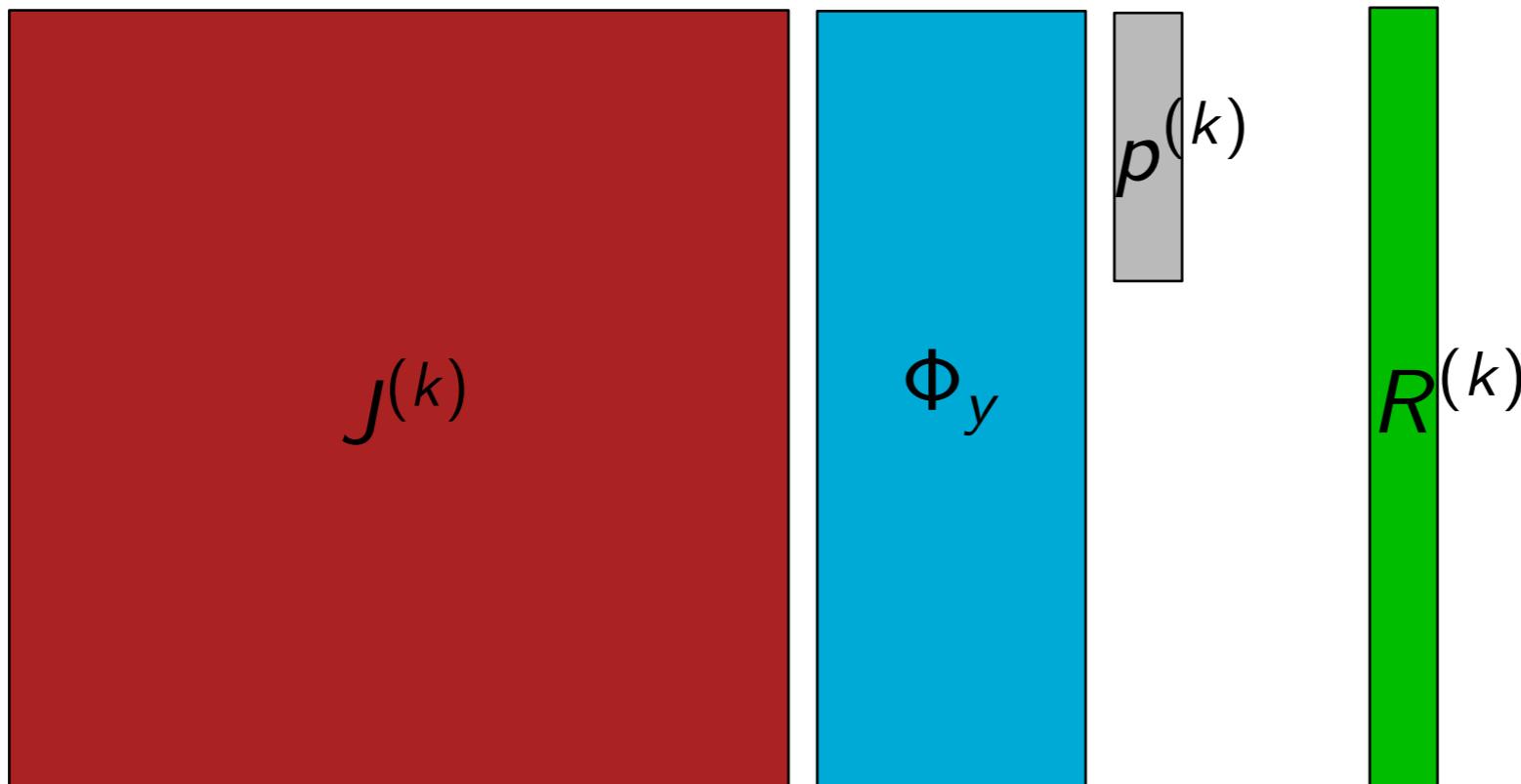
GNAT strategy



Small dimension \neq small cost

- Minimum-residual ROM iterations are

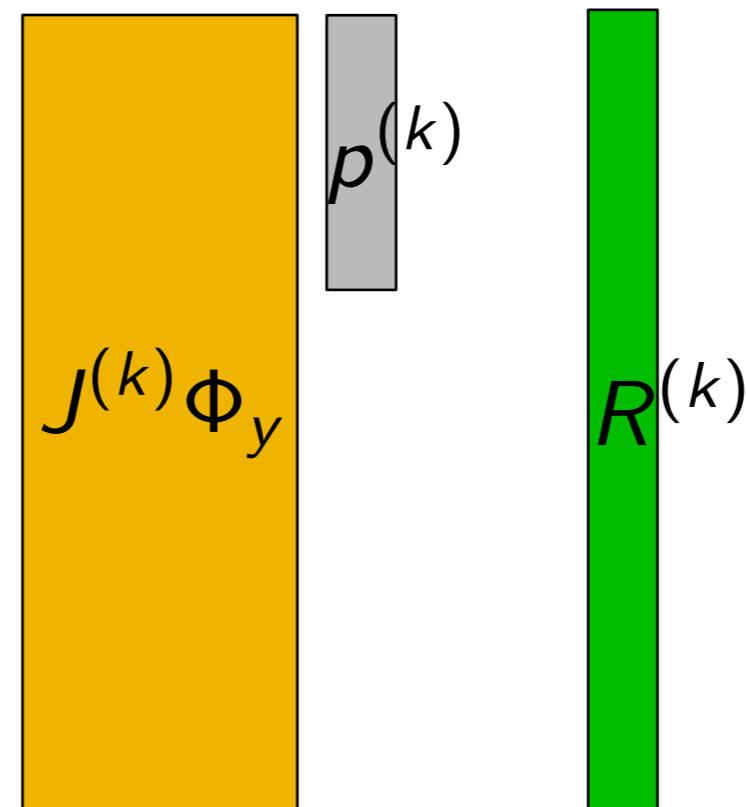
$$\underset{p^{(k)}}{\text{minimize}} \| J^{(k)} \Phi_y p^{(k)} + R^{(k)} \|_2$$



Small dimension \neq small cost

- Minimum-residual ROM iterations are

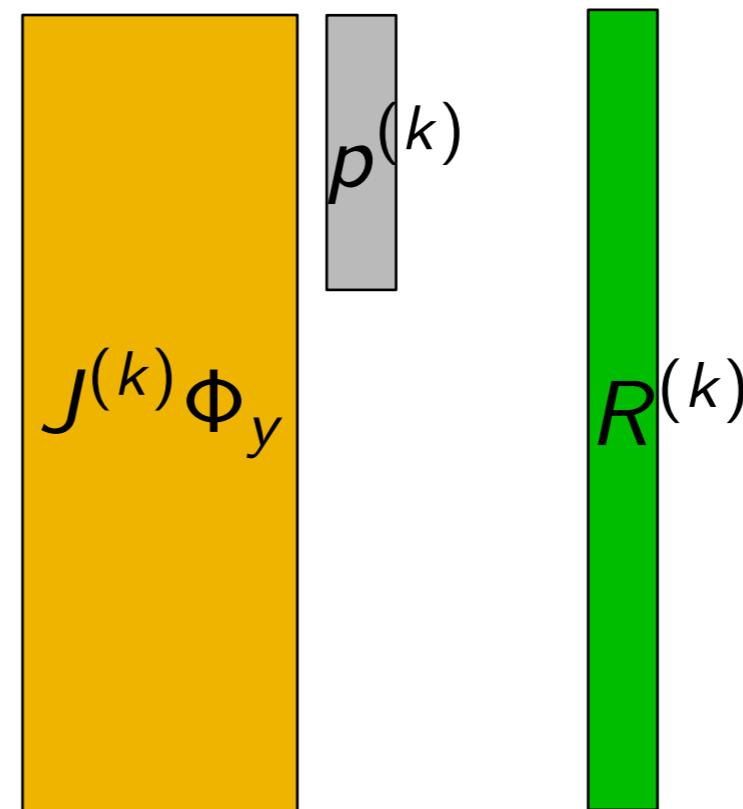
$$\underset{p^{(k)}}{\text{minimize}} \| J^{(k)} \Phi_y p^{(k)} + R^{(k)} \|_2$$



Small dimension \neq small cost

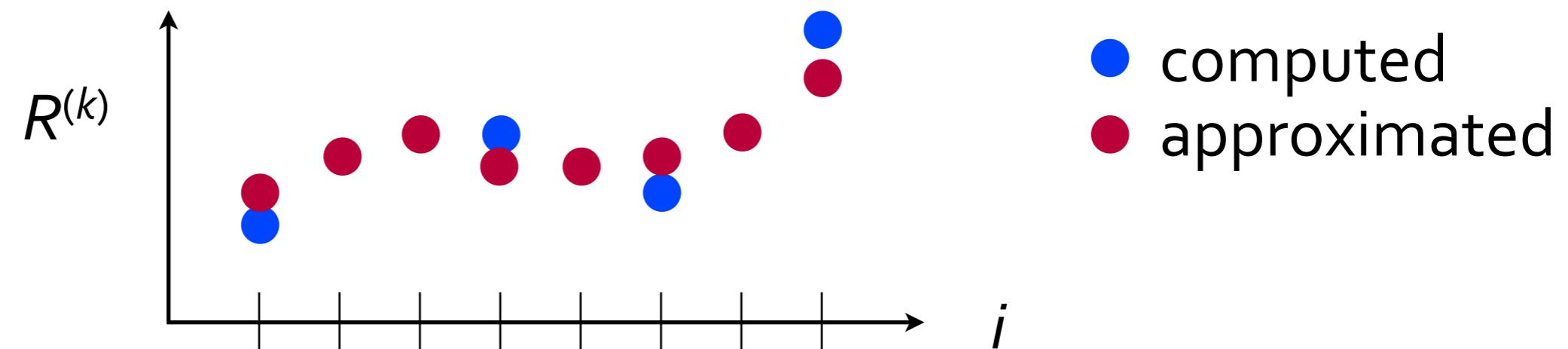
- Minimum-residual ROM iterations are

$$\underset{p^{(k)}}{\text{minimize}} \| J^{(k)} \Phi_y p^{(k)} + R^{(k)} \|_2$$

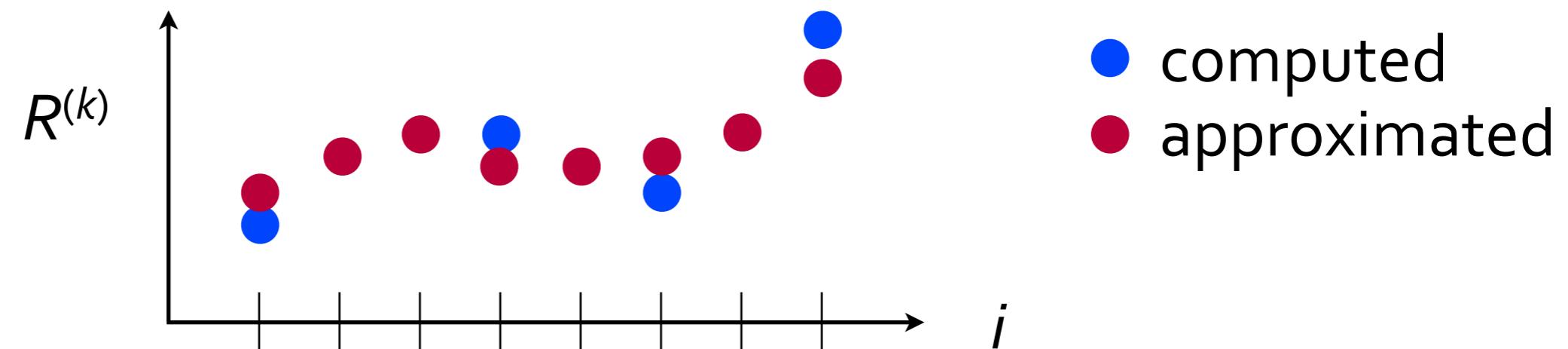


- Operation count scales with N

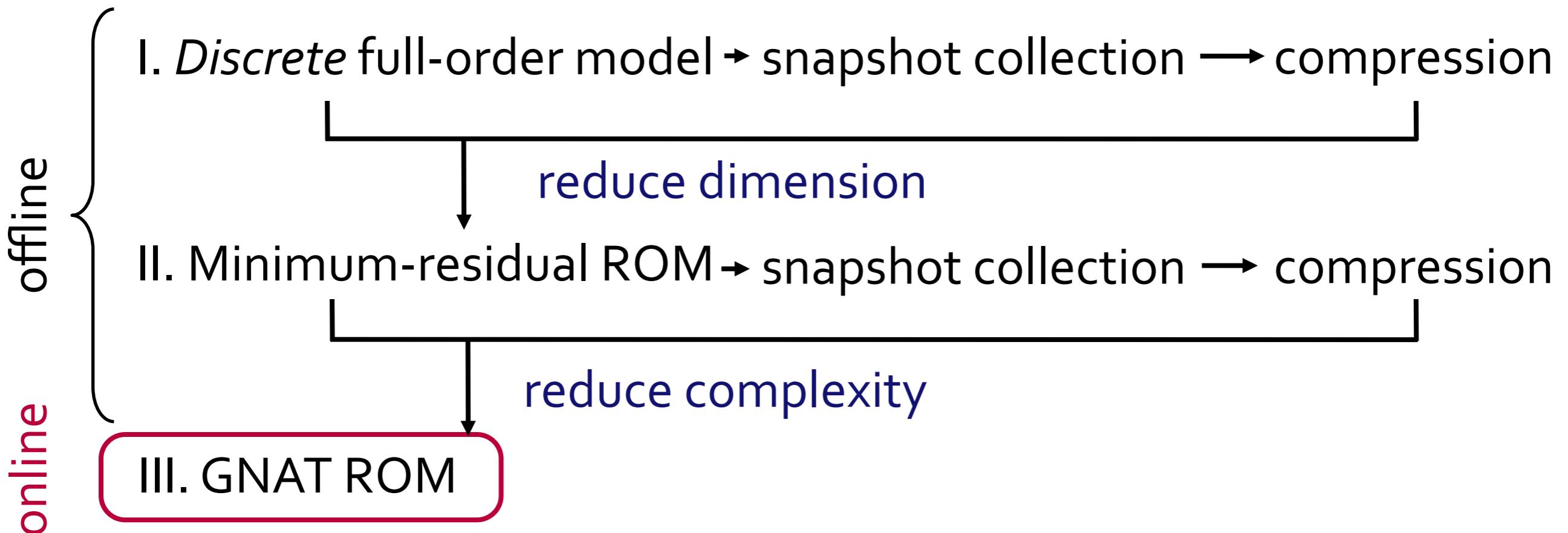
- Goal: approximate the vector $R^{(k)}$
- Given: i) some computed entries of $R^{(k)}$
ii) basis Φ_R
- Procedure: minimize least-squares error at *computed entries*



- Goal: approximate the vector $R^{(k)}$
- Given: i) some computed entries of $R^{(k)}$
ii) basis Φ_R
- Procedure: minimize least-squares error at *computed entries*



- Also apply to each column of $J^{(k)}\Phi$
 - + Discrete optimality: least-squares fit of discrete nonlinear functions
 - + Inexpensive: compute **only a few entries** of $R^{(k)}$



GNAT = discrete-residual minimization + Gappy POD approximation

Outline

- Introduction
- Gauss–Newton with Approximated Tensors (GNAT)
 - Approach
 - Error bound
- Sample mesh
- Example

Error bound

- Assumptions:

1. backward Euler: $R^n(y^n; u) = y^n - y^{n-1} - \Delta t f(y^n; t^n, u)$
2. inverse Lipschitz continuity for $G : (y; t, u) \mapsto y - \Delta t f(y; t, u)$
3. full-order simulation convergence criterion: $\|R^n(y^n; u)\| \leq \epsilon$

Proposition [Carlberg et al., 2013]

The global error at time step n for **any sequence** $(y^0, \tilde{y}^1, \dots, \tilde{y}^{n_t})$ is

$$\|y^n - \tilde{y}^n\| \leq \sum_{k=0}^{n-1} a^k b_{n-k} \leq \sum_{k=0}^{n-1} a^k c_{n-k}$$

Proposition [Carlberg et al., 2012]

The global error at time step n for **any sequence** $(y^0, \tilde{y}^1, \dots, \tilde{y}^{n_t})$ is

$$\|y^n - \tilde{y}^n\| \leq \sum_{k=0}^{n-1} a^k b_{n-k} \leq \sum_{k=0}^{n-1} a^k c_{n-k}$$

- a : inverse Lipschitz constant ‣ \tilde{R} : residual for sequence $\{\tilde{y}^n\}$
- $b_n = \epsilon + \|\tilde{R}^n(\tilde{y}^n; u)\|$ ‣ P : Gappy POD operator
- $c_n = \epsilon + \|P\tilde{R}^n(\tilde{y}^n; u)\| + \|(I - P)\tilde{R}^n(\tilde{y}^n; u)\|$

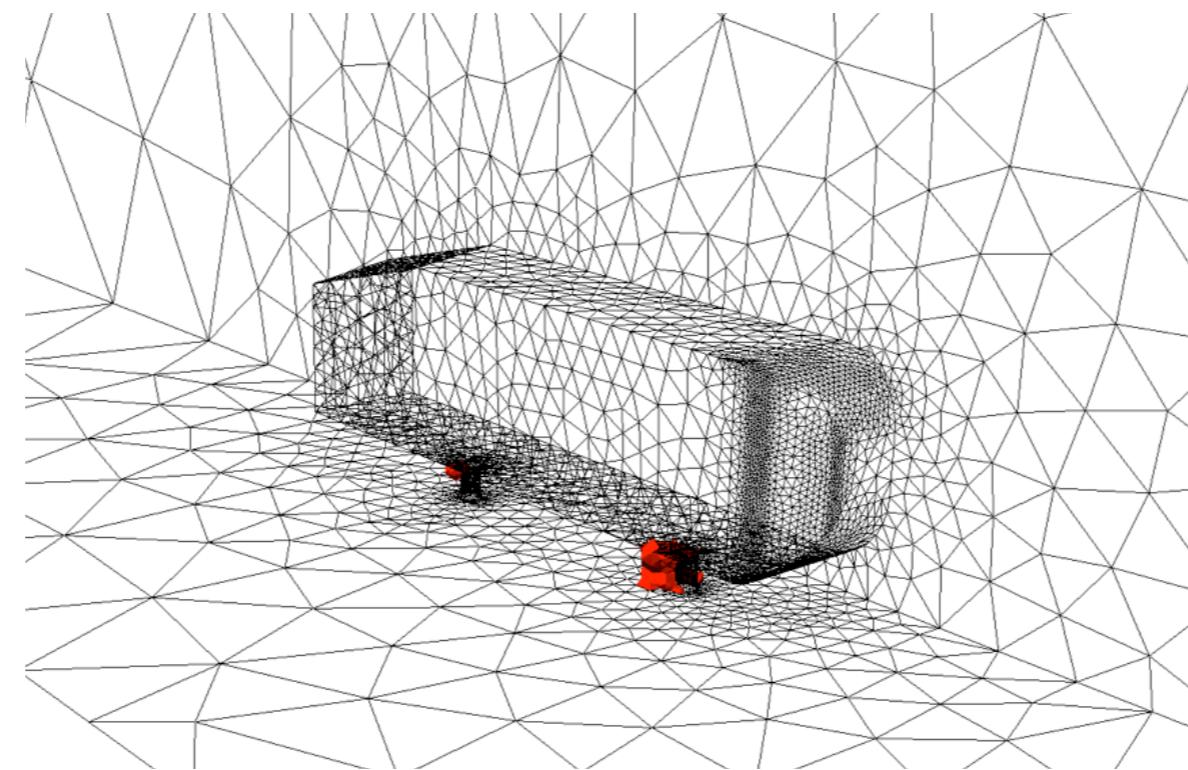
- + Minimum-residual ROM solutions minimize b_n
- + GNAT ROM solutions minimize $\|P\tilde{R}^n(\tilde{y}^n; u)\|$
- + sampling algorithm: heuristic for minimizing $\|(I - P)\tilde{R}^n(\tilde{y}^n; u)\|$
 - discrete optimality enables this!

Outline

- Introduction
- Gauss–Newton with Approximated Tensors (GNAT)
- Sample mesh
- Example

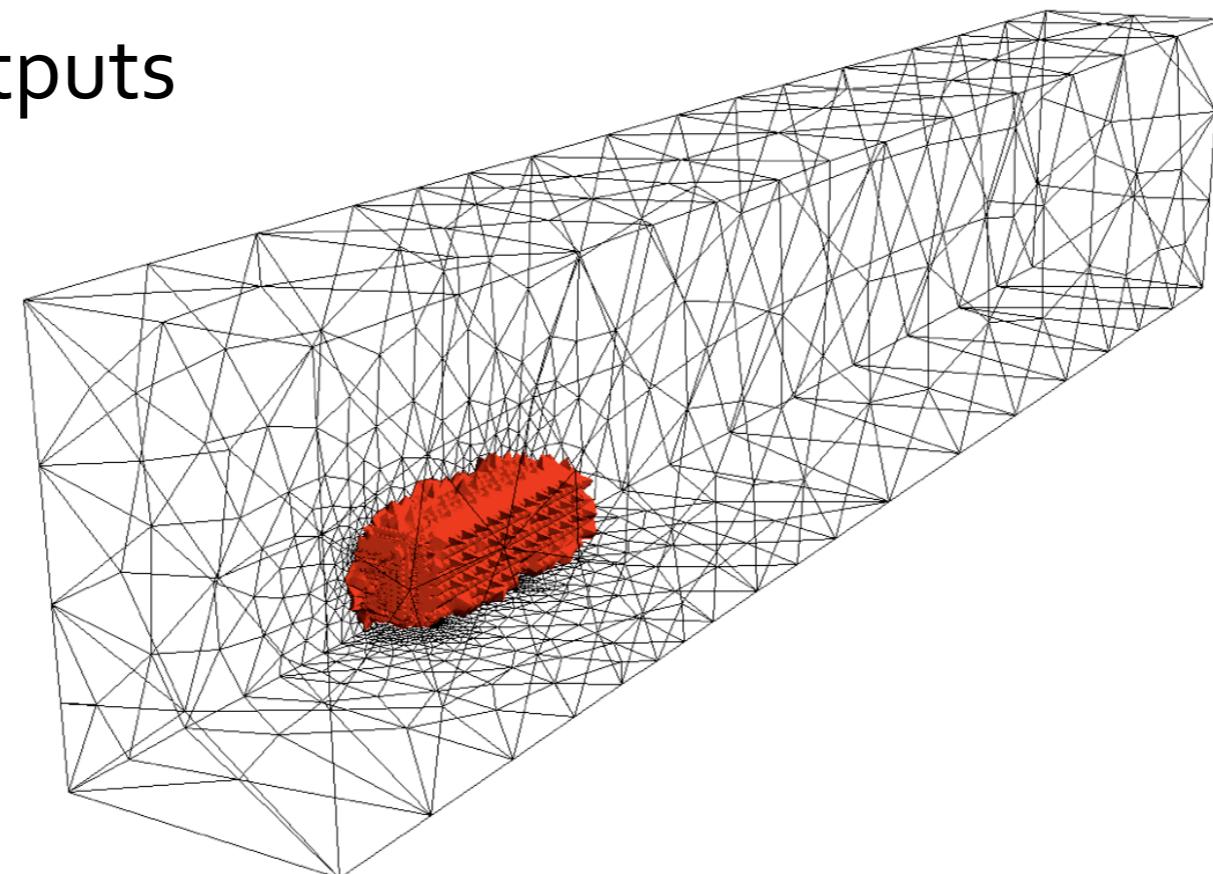
Sample mesh implementation

- *Goals:*
 - reuse existing simulation codes
 - minimize computing cores
 - scalability
- *Key:* GNAT samples only **a few** entries of the residual
- *Idea:* extract minimal subset of mesh



Computing outputs cheaply

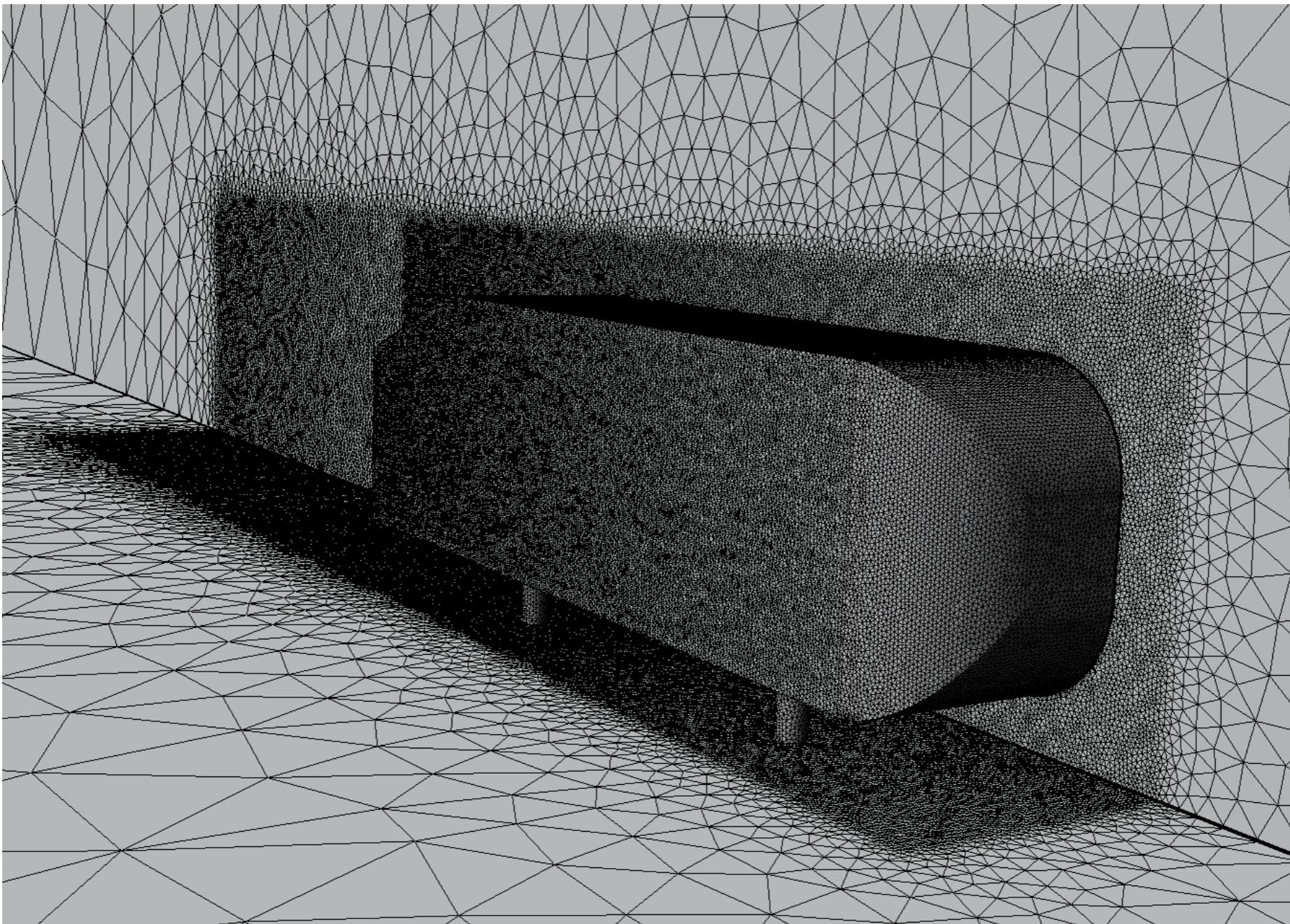
- *Observations:*
 - outputs are often defined locally in space (e.g., lift, drag)
 - outputs may not be computable on sample mesh
- *Output-computation step:*
 1. read POD coefficients y_r computed by GNAT
 2. assemble solution on minimal **output-computation mesh**
 3. compute outputs



Outline

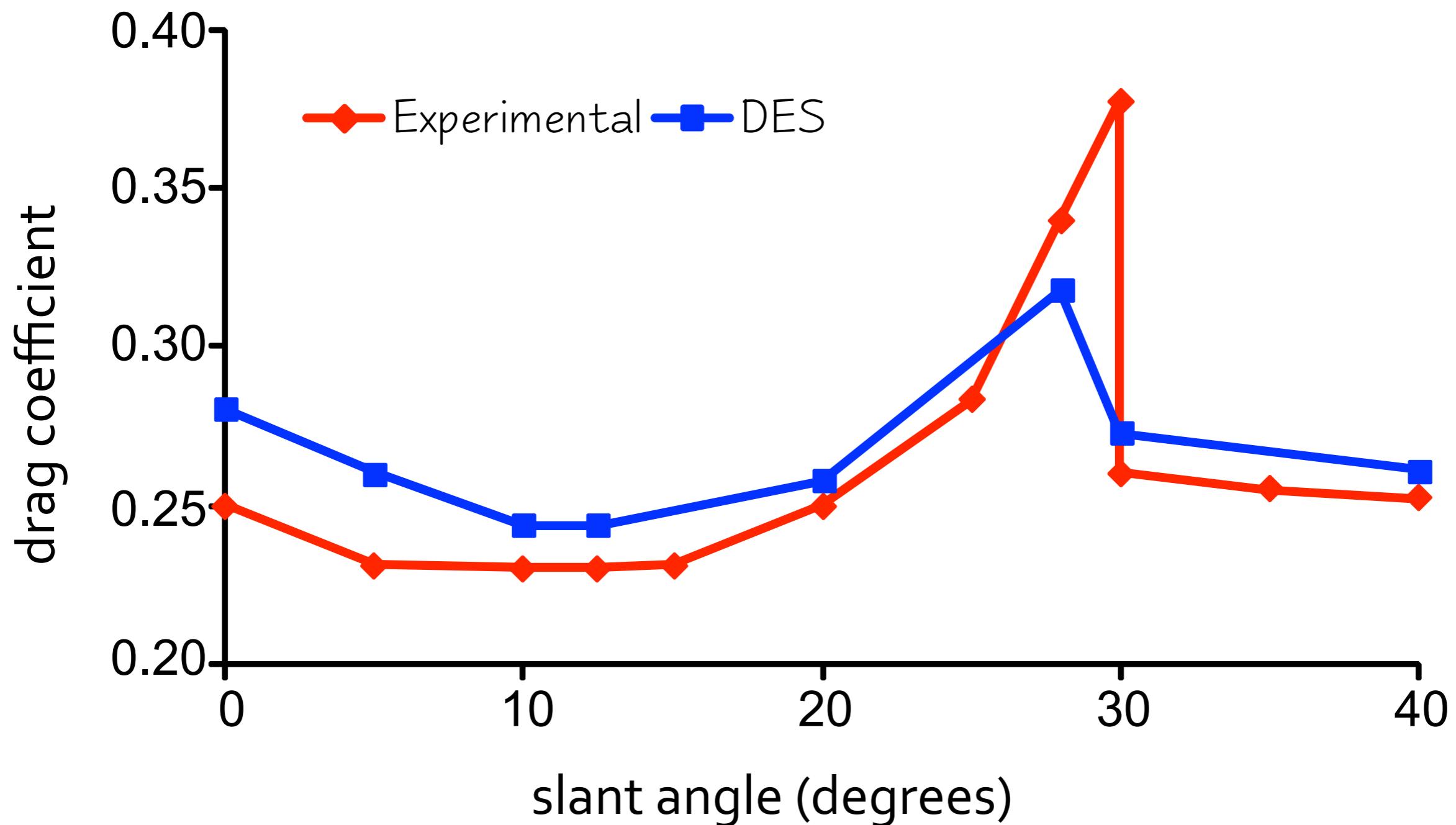
- Introduction
- Gauss–Newton with Approximated Tensors (GNAT)
- Sample mesh
- Example

Example: Ahmed body (fine mesh)

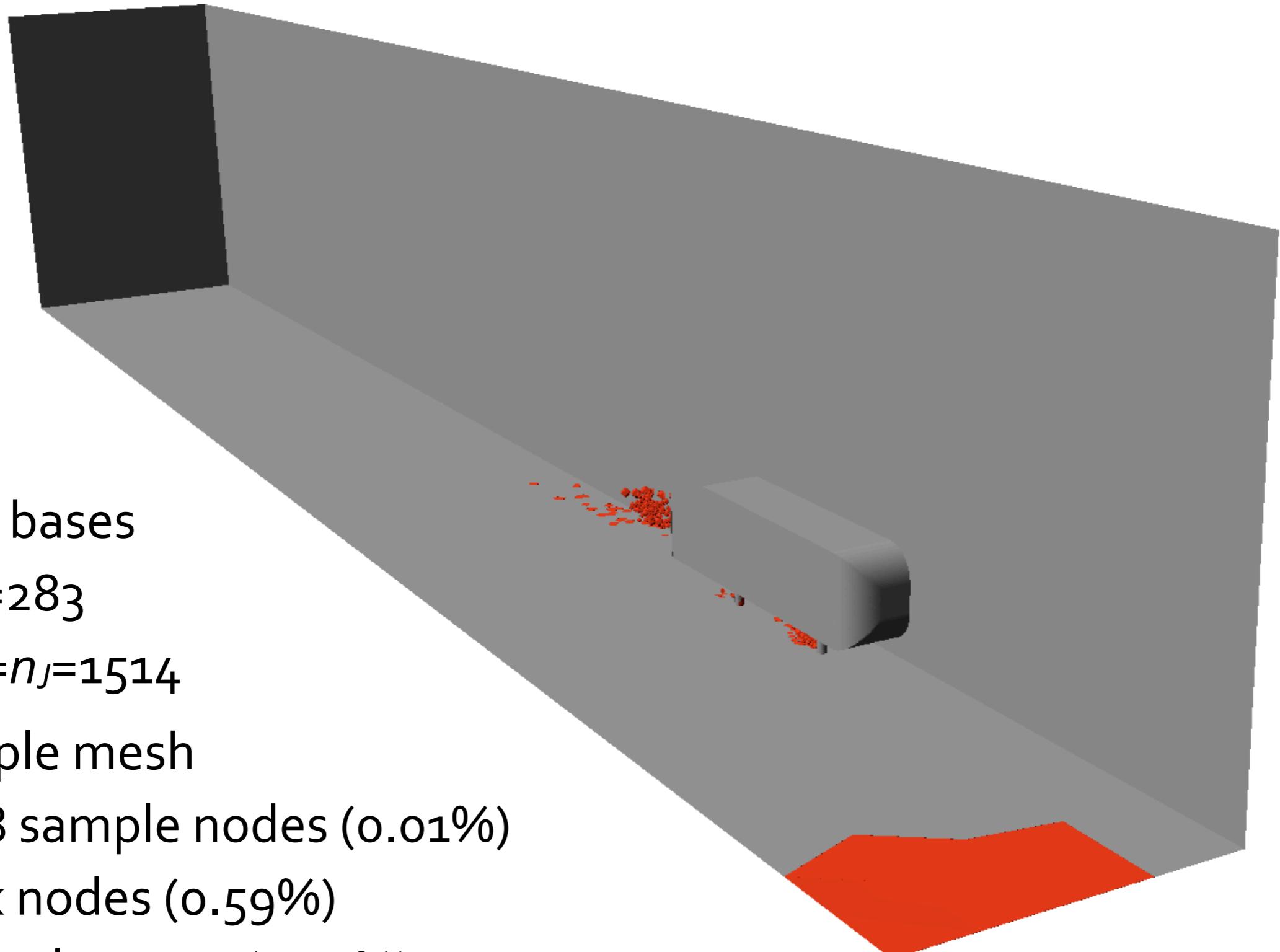


- 2.89×10^6 nodes, 1.70×10^7 tetrahedral volumes
- 1.73×10^7 degrees of freedom

CFD model matches experiment

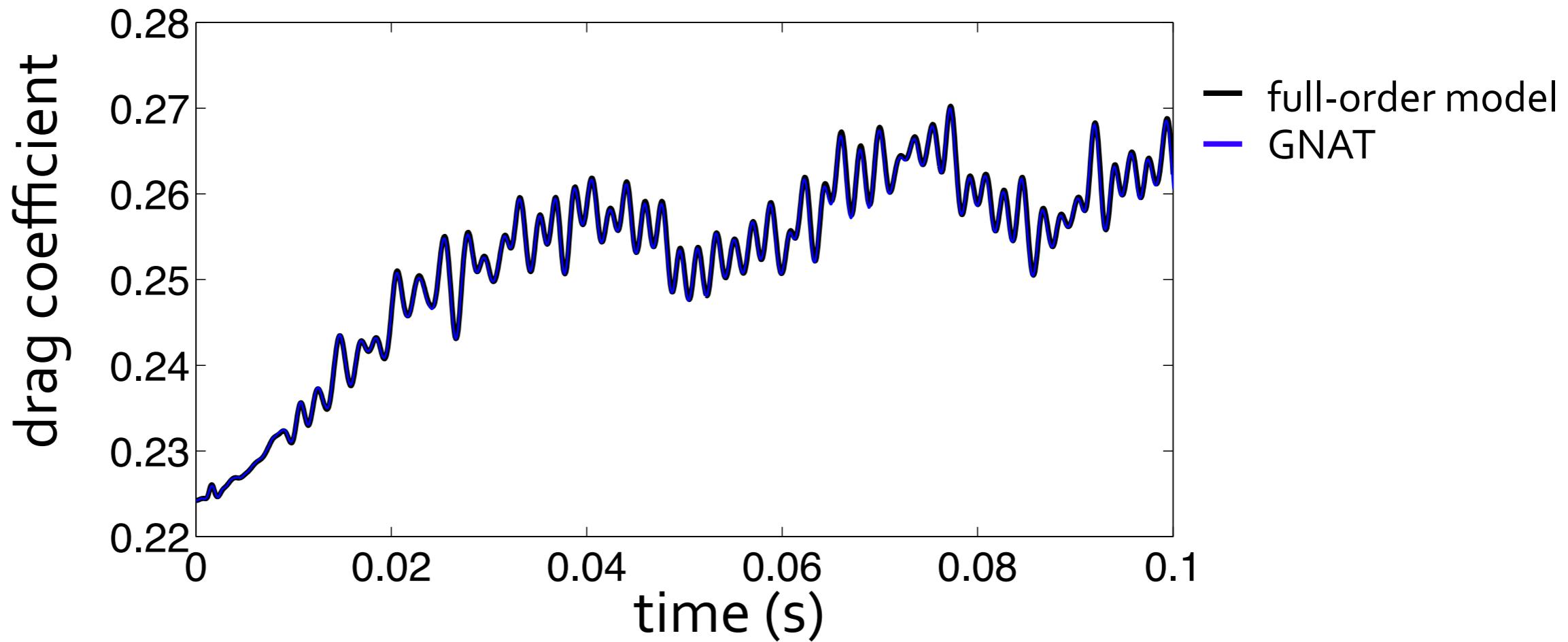


GNAT model



- POD bases
 - $n_y=283$
 - $n_R=n_J=1514$
- Sample mesh
 - 378 sample nodes (0.01%)
 - 17k nodes (0.59%)
 - 56k elements (0.33%)

GNAT results: accurate and fast

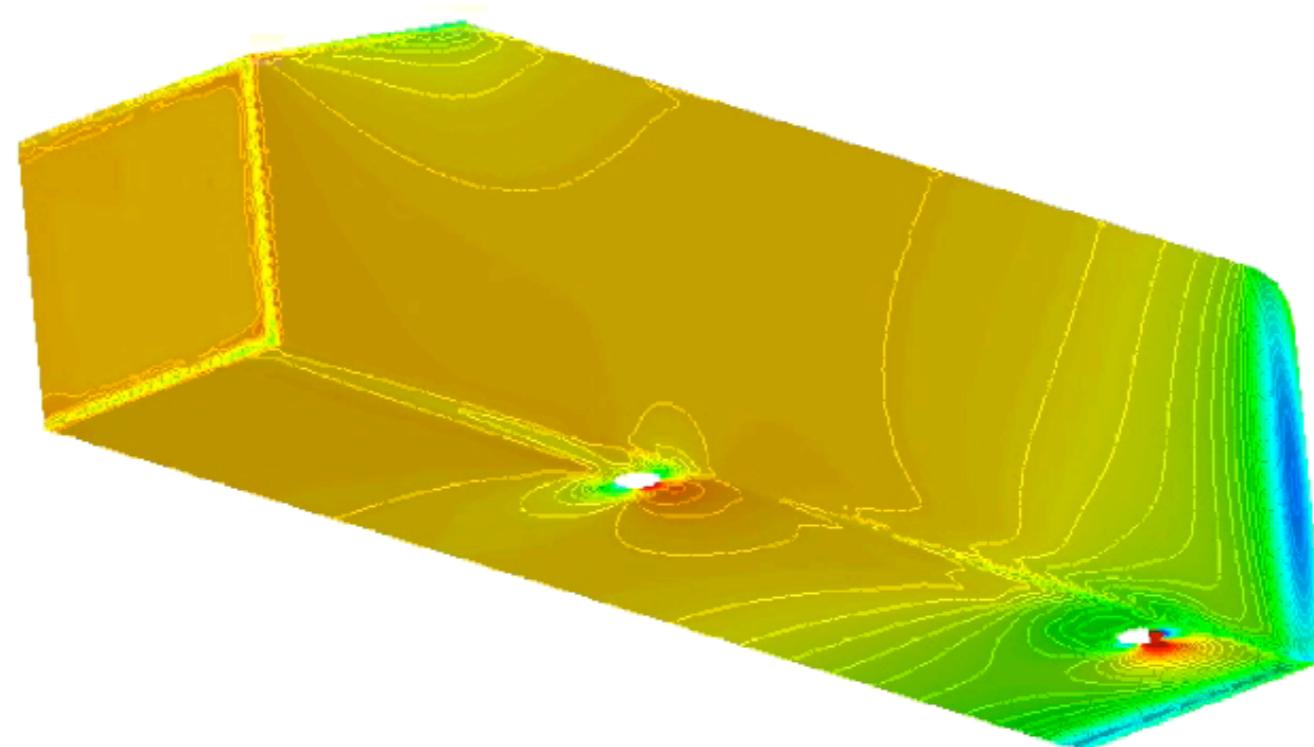


model	relative error	# cores	time, hours	speedup in cpu resources
FOM	-	512	13.3	-
GNAT	0.68%	4	3.88	438

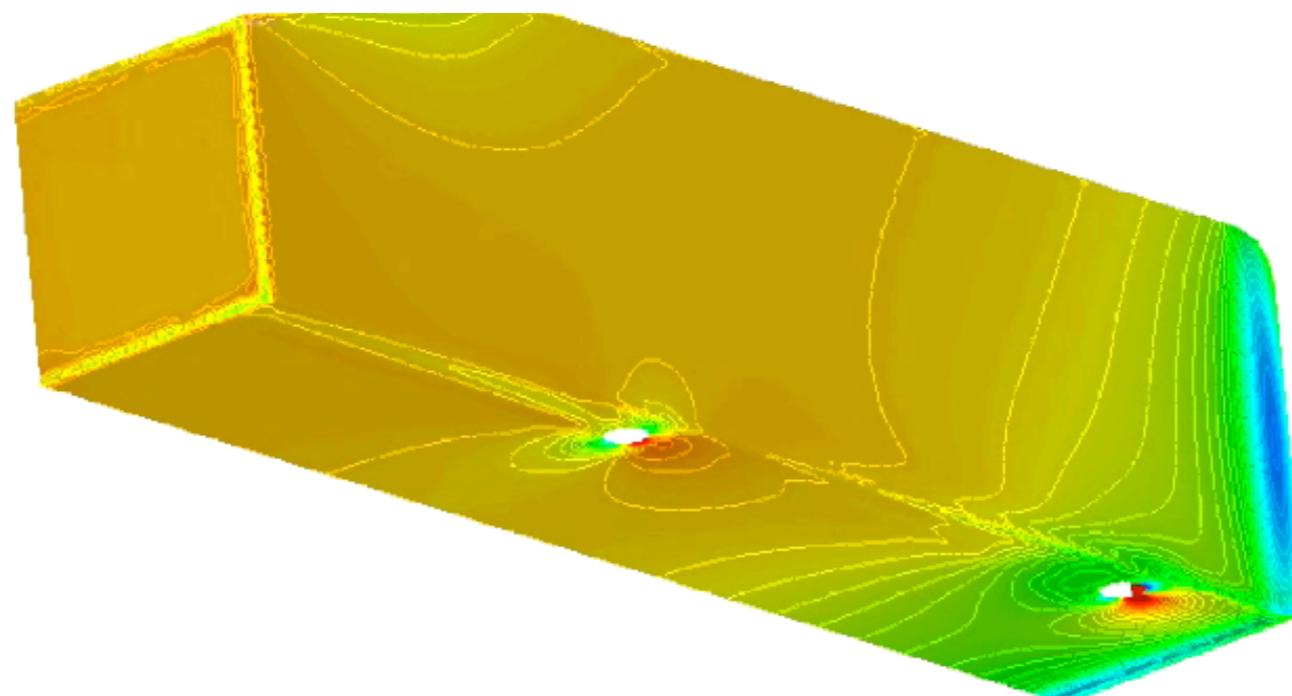
+ negligible error + wall-time decrease + supercomputer → desktop

GNAT results: accurate pressure contours

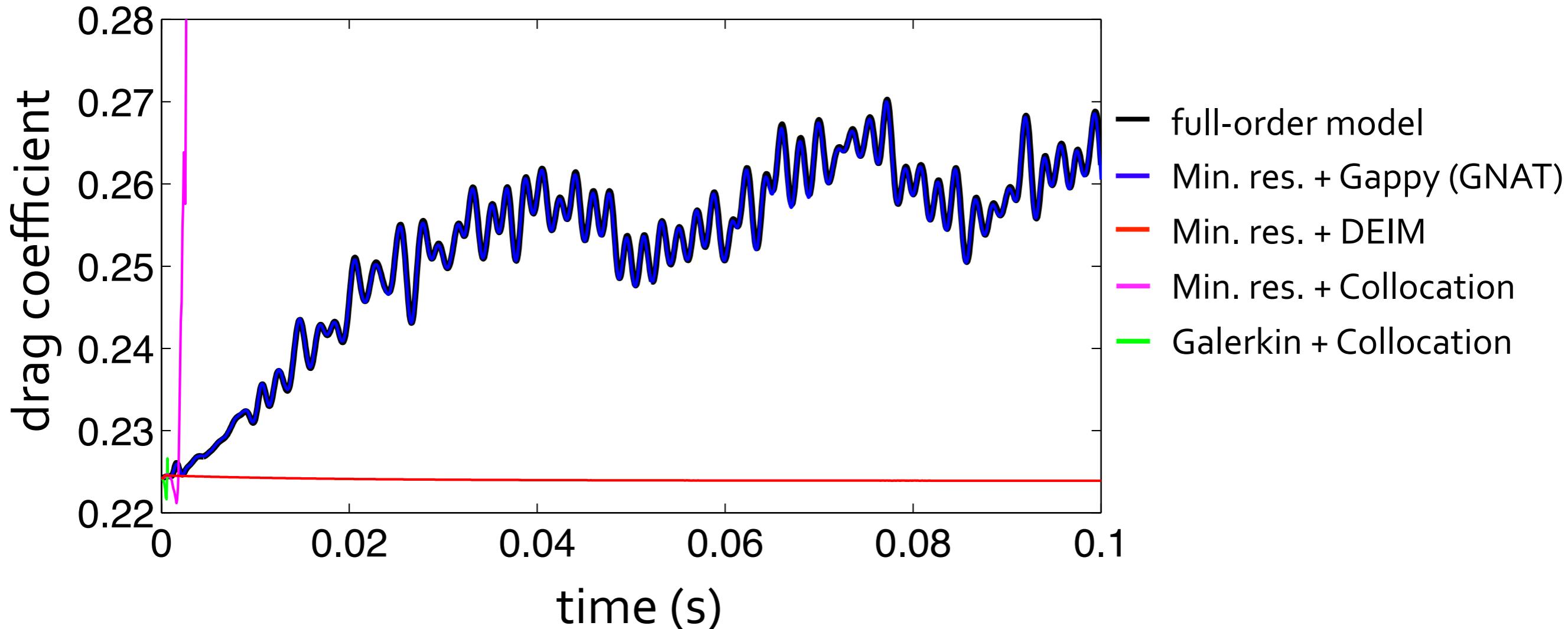
full-order
model



GNAT



- Fixed POD basis and sample mesh (378 sample nodes)



- Galerkin ROM fails
 - discrete optimality important
- Gappy POD: only complexity-reduction method that works!

Summary

- GNAT method
 - discrete-optimal approximations
 - error bound justifies its design
- Sample mesh concept enables many fewer cores
 - supercomputer → desktop
- Ahmed body example
 - speedups over 400, error less than 1%
 - other model-reduction methods failed
- Key papers
 - K. Carlberg, C. Bou-Mosleh, and C. Farhat. "Efficient non-linear model reduction via a least-squares Petrov–Galerkin projection and compressive tensor approximations," International Journal for Numerical Methods in Engineering, Vol. 86, No. 2, p. 155–181 (2011).
 - K. Carlberg, C. Farhat, J. Cortial, and D. Amsallem, "The GNAT method for nonlinear model reduction: Effective implementation and application to computational fluid dynamics and turbulent flows," Journal of Computational Physics, in press, doi:10.1016/j.jcp.2013.02.028 (2013).

1. Preserve classical Hamiltonian/Lagrangian structure [Tuminaro, Boggs]

- *Goal:* preserve key properties (e.g., energy conservation)
- Structural dynamics, molecular dynamics

2. Decrease temporal complexity via forecasting [Ray, van Bloemen Waanders]

- *Goal:* decrease wall-time for nonlinear ROM simulation
- Exploit temporal data to reduce total # Newton its

3. Integrate ROMs within a UQ framework [Sargsyan, Drohmann]

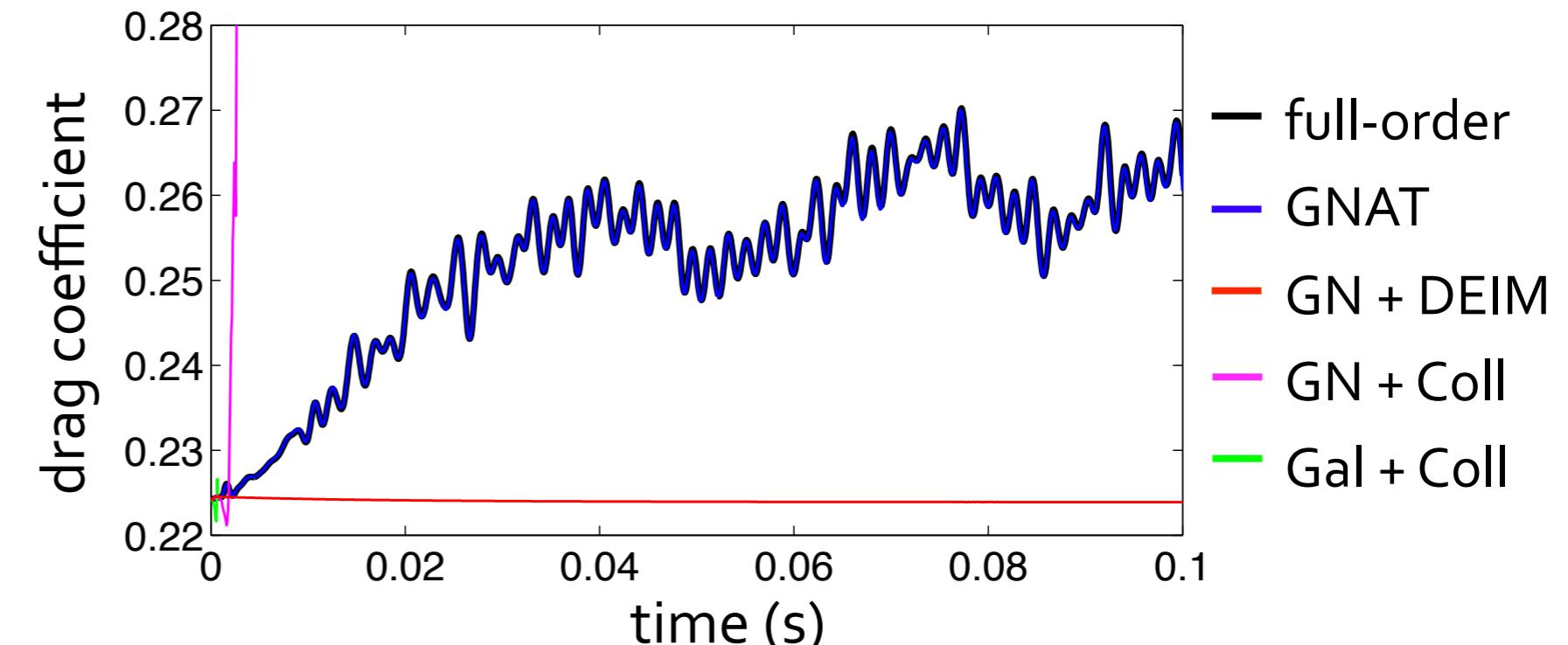
- *Goal:* quantify epistemic uncertainty due to ROM
- Error bounds (i.e., ‘certification’) not useful in UQ

4. ROM interface for Sandia’s simulation codes [Cortial]

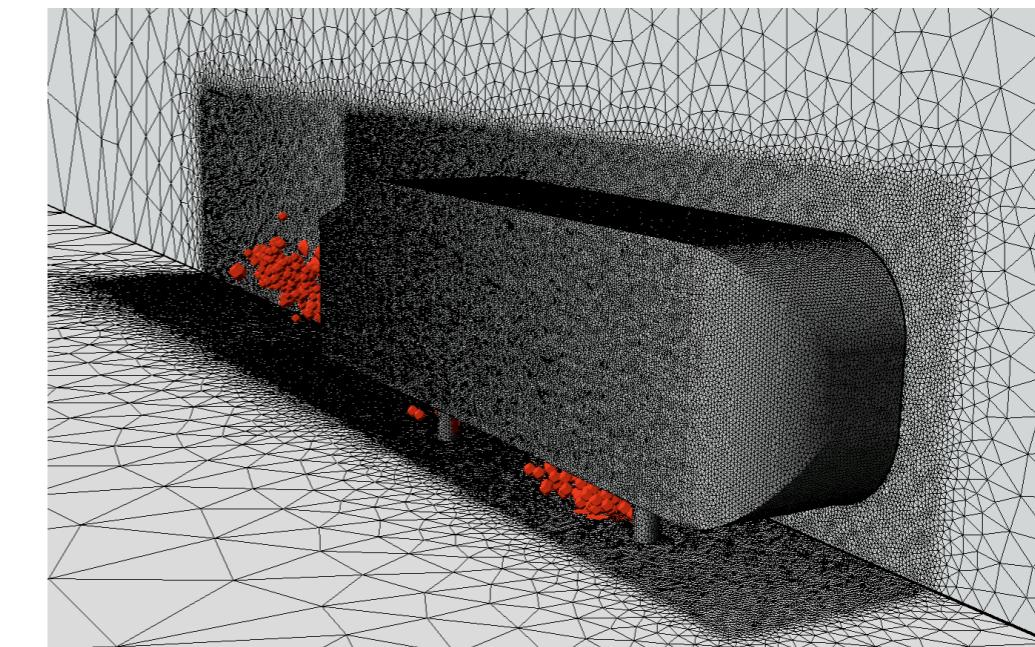
- *Goal:* easily apply nonlinear ROM methods to codes
- non-intrusive Trilinos-based ROM module

Questions?

- Key collaborators
 - Charbel Farhat
 - Julien Cortial
 - David Amsallem



- Funding
 - NSF Graduate Fellowship
 - Toyota Motor Corporation
 - Army Research Laboratory
 - Truman Fellowship program*



* Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under contract DE-AC04-94-AL85000.